


```

/notes="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
E-Coli DH10B"
/db_xref="taxon:9606"
/clone="Plate=2041 Col=13 Row=D"
/clone_lib="CIR Approved Human Genomic Sperm Library D"
/sex="male"
BASE COUNT      60 a      105 c      71 g      104 t
ORIGIN

```

Query Match	9.1%;	Score 221;	DB 40;	Length 340;
Best Local Similarity	86.0%;	Pred. No. 1,49e-265;		
Matches 257;	Conservative	0;	Mismatches 42;	Indels 0;
			Gaps	0

Db	22	TTTTTTTTTTTTTTTGAAGAGATTTCCACTCTGTTGGCCAGAGGCGAGATGCATATGGC	81
Cp	2442	ttttttttttttttttttgtgacagatttctgtctgttaaccagctgtagagtgcaatgyc	23
Db	82	ACGATCTGGGCTCACCAACCTCCGCTCCAGAGTTCAATGATTTCTCTGCTCAGCC	14
Cp	2382	gcmtctcgcgcacgcacactcgcctccccagttcaagnathttttctgcctgagcc	23
Db	142	TCCCGATCTACCTGGATTACAGCACTGTCACCAACGCGCCAGCTAAATTTGATTTTGG	20
Cp	2322	tctgtatgacccagntttccagcaltgcgtccaccagcccggtctaatttgtatttlnag	22
Db	202	TAGAGATGGGGTTCTCATGTTGGTCAAGCGTGTCTTGAGACTCCAACTCAGTGATC	26
Cp	2262	tagagatgaggttttccaccglttgcagctgctgttctaactccgcgcttagatgaltc	22
Db	262	CACCCACTCAGGCTCCCAAAATGCTGGGATTAACAGGATAGAGCAACGGGCTGGCCA	320
Cp	2202	caccaccactcgcctcccaaatgtctggaattaccaagtgtgagaccactcgccttgcca	214

RESULT	2	389 bp	DNA	GSS	23-MAR-1999
LOCUS	A0429198				
DEFINITION	CITR1-EL-2581D3.1.FE CITR1-EL Homo sapiens genomic clone 2581D3 genomic survey sequence.				
ACCESSION	A0429198				
NTID	94496964				
VERSION	A0429198.1	GI:4496964			
KEYWORDS	GSS.				
SOURCE	human.				

Eukaryota; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 389)
 Zhao, S., Adams, M.D., Niernan, W., Malek, J., Shizuya, H., Simon, M. and
 Venter, J. C.
 Title
 Use of BAC End Sequences from Caltech Libraries for Sequence-Ready
 Map Building
 Journal
 Unpublished (1997)

Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: bhee@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/cdb/humgen/bac_end_search/bac_end_search.html
Seq primer: M13-21
Class: BAC ends.

```

FEATURES
source      location/Qualifiers
1..389
/organism="Homo sapiens"
/ncbi="Vector: pBelOAC11, Site_1: EcoRI, Site_2: EcoRI"
Caltech Human BAC Library D"
/db_xref="taxon:9606"
/clone="2581d3"
/clone_1bp="CITB1-E1"

```

BASE COUNT	74 a	114 c	86 g	115 t
ORIGIN				

Query Match	9.1%;	Score 222;	DB 42;	Length 389;
Best Local Similarity	85.5%;	Pred. No. 4.98e-268;		
Matches	260;	Conservative	0;	Mismatches 44;
			Indels	0;
			Gaps	0;

Db	29	TTTTTCTTTTTTGAGACAGAACTTTTGGCTCTGTGTGTCGCCACAGCGTGAGTGCAAAGCGC	88
Cp	2440	tttttttttttttttgacagagtcttgcctctgtttaccagcgctggagtgaaatgagcg	2384
Db	89	CATCTACGCTCAGCGAATCTCCGCGCTCCCGGGTTTAAACAATTCCTCGCTCGGGCTC	148
Cp	2380	gntctaggcttcaacgcaactctcgccctcccaattcaaaatunttctgcctcgccctc	2321
Db	149	CTGAGTAGCTGGGATTACAGCGCATGCACACACGCGCTGGCTAATTTTGTATTTTAACTA	208
Cp	2320	ctgagtagccagagntttcttcacgcaatgcgtacacacgcccgcctaaattngtattctnagta	2261
Db	209	CAGACGGGGTTTCCATGTTGGTCAGAGCGGAGCTGTGAATCCGACCCCTAATGATATCA	268
Cp	2260	gagatggaggtttctccaccgtttgcacggctgctgtcttgaactccggctcttagatgtcca	2201
Db	269	CTTACTTCAGCCTCCGAGATGCTAGGATTACAGCGGTGAGCCACGCGCCTTGCGGCTGT	328
Cp	2200	ccccacttgcgctccccaanaagttgttgatattacagagtgtgtagcacatctgcgcttgcacaga	2141
Db	329	CTTG 332	
Cp	2140	cttg 2137	

RESULT	3			
LOCUS	A0372425	556 bp	DNA	GSS
DEFINITION	RPC111-143N11.TV RPC111 Homo sapiens genomic clone R-143N11, genomic survey sequence.			
ACCESSION	A0372425			
NTID	04343448			
VERSION	A0372425.1	GI:4343448		
KEYWORDS	GSS.			
SOURCE	human.			

REFERENCE
AUTHORS
TITLE
JOURNAL

Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbeet@igr.org
For clone availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
Research Genetics (inforesgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/Bac_end_search/Bac_end_search.html
Seq primer: T7
Class: BAC ends.

```
FEATURES
Source
location/Qualifiers
1..556
/organism="Homo sapiens"
/note="Vector: pBAC3.6; Site.1: EcoRI; Site.2: EcoRI
RcCl1 Human Male BAC Library"
/db_xref="GBD:7554850"
/db_xref="taxon:9606"
```


Query Match	8.8%	Score 214	DB 14	Length 325
Best Local Similarity	85.6%	Pred. No. 3	16e-256	
Matches	256	Conservative	0	Mismatches 42; Indels 1; Gaps 1
Db	5	TTTCCTTTTTTTTTTTGAGCGGAGTTTCGCTCTTGTATACCGAGCTGAGTCAGTCGC	64	
Cp	2442	tt	3283	
Db	65	GTGATCTCGGCTCACCCGACACCTCTGCCTCCTCGGTTACCAATTCCTCGCTCAGAC	124	
Cp	2382	gcgcttcggggtcacccgaacacctgcctcccccagcttaagaaatcttcgcctcggcc	2323	
Db	125	TCCGAGTAGCTGGGATTACAGACATGTGCATCAGCCCGCTCAATTTTATATTTTA	184	
Cp	2322	ttctgagtagcagattacagtcacatgctacacacgcgcggctaatlttg-tattttna	2264	
Db	185	GTAAGACGCGGGTTTCTCCATGTTGGTTAGGCTGCTGTCCAATCCGACCTCAGGTAT	244	
Cp	2263	gtagagatggggtttcttcacagttgtltaaggctggtcttaactccggctttagatg	2204	
Db	245	CGGCCACCTCGGCGCTCCCAAGTCTGTGGGATTACAGCAATGAGGCACCGACCGCGC	303	
Cp	2203	ccaccaccttgcctccccaagtgctggagtttaagagtgtagacacctgcgctcgcc	2145	
RESULT	6	A1572680	352 bp	mrna
LOCUS				EST
DEFINITION		tr78a09_x1 NCI CGAP Panl Homo sapiens cDNA clone IMAGE:2224408	3'	29-MAR-1999
		similar to contains Alu repetitive element,, mRNA sequence.		
ACCESSION		A1572680		
NID		94536054		
VERSION		A1572680.1		GI:4536054
KEYWORDS		EST.		
SOURCE		human.		
ORGANISM		Homo sapiens		
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;		
		Eutheria; Primates; Catarrhini; Homiidae; Homo.		
REFERENCE		1 (bases 1 to 352)		
AUTHORS		NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .		
TITLE		National Cancer Institute, Cancer Genome Anatomy Project (CGAP),		
		Tumor Gene Index		
JOURNAL		Unpublished (1997)		
COMMENT		On May 18, 1998 this sequence version replaced gi:3136689.		
		Contact: Robert Strausberg, Ph.D.		
		Tel: (301) 496-1550		
		Email: Robert.Strausberg@nih.gov		
		Life Technologies catalog #: 11548-013		
		DNA Sequencing by: Washington University Genome Sequencing Center		
		Clone distribution: NCI-CGAP clone distribution information can be		
		found through the I.M.A.G.E. Consortium/LNL at:		
		www-bio.linnl.gov/bdrip/image/image.html		
FEATURES		Seg primer: -40UP from Glibo		
		High quality sequence stop: 309.		
SOURCE		Location/Qualifiers		
		1..352		
		/organism="Homo sapiens"		
		/note="Organ: pancreas; Vector: pCMV-SPORT6; site:1: SalI;		
		site:2: NotI; Cloned unidirectionally. Primer: Oligo dT.		
		Average insert size 1.72 kb. Life Technologies catalog #:		
		11548-013"		
		/db_xref="taxon:9606"		
		/clone="IMAGE:2224408"		
		/clone_11b="NCI CGAP_Panl"		
		/tissue_type="adenocarcinoma"		
		/lab_host="DH10B"		
BASE COUNT		72 a 102 c 76 g 102 t		
ORIGIN				
Query Match	8.8%	Score 214	DB 29	Length 352
Best Local Similarity	84.9%	Pred. No. 3	16e-256	
Matches	253	Conservative	0	Mismatches 45; Indels 0; Gaps 0

Query Match	8.8%	Score 215	DB 15	Length 421
Db 1	TTTTTTTTTTTTTTGAGCGAGGATTCACTCTGTGTTGCCAGGCTGAGTGCAGTAAGGCGTG	60		
Cp 2439	tt	2380		
Db 61	ATCTTGACTCACTGCACCTCTGCTCTCTGGGTTCCAGCAATTCCTCTGCTCAGCCTCC	120		
Cp 2379	ntcttggtctacacgaacctctgcctcccccagattcaagaaatttctgcctcgcgtccc	2320		
Db 121	CGAGTAGCTGGGATTACAGGCGATGTGCTACCATGCTTGCTGGCTAAATTTTGTATTTTTAG	180		
Cp 2319	tgaagtacgaagnttaccagcatgctgcaccagcgcgcgtcaatttgattttagatag	2260		
Db 181	AGATGGGCTTCTCATCTTGTGTCAGGCTGTGATATCTCAACATCCCAACTCAGATGATCCAC	240		
Cp 2259	agatggggtttctctcacgttgctcaggtgtgtcttgtaactccggccttagatgtaccac	2200		
Db 241	CCACCTCGGCGCTCCCAAGTCTGGGATTACAGCGATGAGCCACACACCGACCCGAC	238		
Cp 2199	ccaccttgccctcccaagtggttgattacaagtggtgagccactgcgcctcggccaag	2142		
RESULT	7			
LOCUS	AA569220	421 bp	mRNA	EST 09-SEP-1997
DEFINITION	nm30h04.s1 NCI-CGAP_Lip2 Homo sapiens CDNA clone IMAGE:1061719			
	similar to contains Alu repetitive element; contains element ITR6			
	repetitive element ;, mRNA sequence.			
ACCESSION	AA569220			
NID	92342274			
VERSION	AA569220.1	GI:2342274		
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;			
	Eutheria; Primates; Catarrhini; Hominoidea; Homo.			
REFERENCE	1 (bases 1 to 421)			
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .			
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),			
	Tumor Gene Index			
JOURNAL	Unpublished (1997)			
COMMENT	On Jan 24, 1995 this sequence version replaced gi:634330.			
	Contact: Robert Strausberg, Ph.D.			
	Tel: (301) 496-1550			
	Email: Robert.Strausberg@nih.gov			
	Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.			
	Emmert-Buck, M.D., Ph.D.			
	CDNA Library Preparation: David B. Krizman, Ph.D.			
	CDNA Library Arrayed by: Greg Lennon, Ph.D.			
	DNA sequencing by: Washington University Genome Sequencing Center			
	Clone distribution: NCI-CGAP clone distribution information can be			
	found through the I.M.A.G.E. Consortium/LMW at:			
	www.bio.liml.gov/db/tp/image/image.html			
FEATURES				
Source	Insert Length: 1484 Std Error: 0.00			
	Seq primer: -40m13 fwd, ET from Amersham.			
	Location/Qualifiers			
	1..421			
	/organism="Homo sapiens"			
	/note="Vector: pAMP10; mRNA made from liposarcoma, CDNA			
	made by oligo-dT priming. Non-directionally cloned.			
	Size-selected on agarose gel, average insert size 600 bp			
	Reference: Krizman et al. (1996) Cancer Research			
	56:5380-5383."			
	/db_xref="taxon:9606"			
	/clone="IMAGE:1061719"			
	/clone_1lb="NCI-CGAP_Lip2"			
	/tissue_type="liposarcoma"			
	/lab_host="DH10B"			
BASE COUNT	125 a 93 c 115 g 88 t			
ORIGIN				

DB	Best Local Similarity	85.2%	Pred No.	1,066-257	Matches	253	Conservative	0	Mismatches	44	Indels	0	Gaps	0
Db	85	GGCCAGCGCAGTGCCTCACGCCCTGTATCCCAACACTTTGGAGGCCGAGGTGGCGGA	144											
Qy	2145	ggccaagcgcaatgagctcacaccttlaaccacaacacttggagagccaaggtgtgtgga	2204											
Db	145	TCACCTGAGGTTGGAGTTTGGACACAGCGCTCACCACATGATGAGAGAAACCCGGTGTAC	204											
Qy	2205	tcctctaaagccggaggttccaagaaccagctctaccaaggttgagaacccatcttca	2264											
Db	205	AAAAATACAGATTTACCGACGCGATGGTGGCGCATGCTGTATCCCACTACTACGAGG	264											
Qy	2265	naaatatcaaatatgacggcggtgtgtagcatgtgtctlaancctgtactactagaag	2324											
Db	265	CTGAGCGACGAGAACCTGCTTGAACCTCAGAGAGTGGAGTGTGACGTGACGACAGATCGCAC	324											
Qy	2325	ccgaagaggaanaatttcttgaacttggagaggaaggttcggtgagcccaagcgc	2384											
Db	325	CATTGCCCTCCACCTCGGTACACAGAGGGAACCTGTCTTCACAGAAAAATAA	381											
Qy	2385	cattgcctccaccccggtgaacaagaggaacacctgtccaaaaaataaaaaa	2441											
RESULT	8	A0093841	432 bp	DNA	GSS	25-AUG-1998								
LOCUS		HS-3023_B2_G10_T7	CIT	Approved Human Genomic Sperm Library D	Homo sapiens genomic clone	Plate=3023 Col=20 Row=N	genomic survey							
DEFINITION														
ACCESSION		A0093841												
NID		93465288												
VERSION		A0093841.1	GI:3465288											
KEYWORDS		GSS.												
SOURCE		human.												
ORGANISM		Homo sapiens												
		Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;												
		Primates; Catarrhini; Hominoidea; Homo.												
REFERENCE		1 (bases 1 to 432)												
AUTHORS		Mahairas,G.G., Zackione,K.D., Smith,T., Tipton,S., Schmidt,S.,												
TITLE		Talcoff,R., Abajian,C., Blanchard,A., West,A. and Hood,L.E.												
		Construction of a Characterized Clone Resource for Genomic												
		Sequencing: Generation and Preliminary Analysis of 20,000 Sequence												
JOURNAL		Tagged Connectors												
COMMENT		Unpublished (1997)												
		Contact: Mahairas GG, Wallace JC, Hood L												
		High Throughput Sequencing Center												
		University of Washington												
		401 Queen Anne Avenue North, Seattle, WA 98109, USA												
		Tel: (206) 616-3618												
		Fax: (206) 616-3887												
		Email: jwallace@u.washington.edu												
		Sequence Tagged Connector												
		Plate: 3023 row: N column: 20												
		Class: BAC ends												
FEATURES		High quality sequence stop: 432.												
		Location/Qualifiers												
SOURCE		1. 432												
		/organism="Homo sapiens"												
		/note="Organ: sperm; Vector: pbeloBAC1; BAC clones in												
		E-Coli DH10B"												
		/db_xref="taxon:9606"												
		/clone="Plate=3023 Col=20 Row=N"												
		/clone_lib="CIT Approved Human Genomic Sperm Library D												

QY	2145	gcccagggcgagtggtccacacccctgtaatcccaacactcttgagggccaaagytgggtgca	2204				
Db	97	TCACCTGAGTGAGGACTTTCGAGACCGCCGACCACATGATGGAACCCCATCTACT	156				
OY	2205	tcatctaaagccggaggttcaagaccgagccgctgaccacagctgagaaaccccatcttaact	2264				
Db	157	AAAAATCAAAATATGACCGCGCGCTGTGTGCGACATCGCTGTAACTCCAGCTACTATGAGG	216				
OY	2265	naaataacnaaattagatgcgcggcgctggttagcgcatgctgtcaancctgctactcaaggag	2324				
Db	217	CTGAGCGAGAGAAATCGCTTGACCTTGAGCGACGAGAGTTCCGCTGAGCCGACATCACGC	276				
OY	2325	ccgagcagaanaattcttcttgaacttgaggagcgagaggttcggttgagcccaagancg	2384				
Db	277	CATTGCACCTCCAGCCTCGGCAACAGAGTGAACCTCCATTCATATAAAAAATATGAAAA	334				
OY	2385	catgtcacctccagcctcggtgtatacaagacaaactctgtccadaaaaaaaaaaaaaa	2442				
RESULT	9	AT567676	454 bp	MRNA	EST	26-MAR-1999	
LOCUS		tr64b05.x1	NCI-CGAP_Pan1	Homo sapiens	CDNA clone	IMAGE:2223057	3'
DEFINITION							
ACCESSION		AF567676					
NID		94526128					
VERSION		AT567676.1					
KEYWORDS		EST.					
SOURCE		human.					
ORGANISM		Homo sapiens					
REFERENCE		Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;					
AUTHORS		Eutheria; Primates; Catarrhini; Hominiidae; Homo.					
TITLE		1 (bases 1 to 454)					
JOURNAL		NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .					
COMMENT		National Cancer Institute, Cancer Genome Anatomy Project (CGAP),					
		Tumor Gene Index					
		Unpublished (1997)					
		On Jun 5, 1998 this sequence version replaced gi:3187781.					
		Contact: Robert Strausberg, Ph.D.					
		Tel: (301) 496-1550					
		Email: Robert.Strausberg@nih.gov					
		Life Technologies catalog #: 11548-013					
		DNA Sequencing by: Washington University Genome Sequencing Center					
		Clone distribution: NCI-CGAP clone distribution information can be					
		found through the I.M.A.G.E. Consortium/LNLW at:					
		www-bio.llnl.gov/dbtrp/image/image.html					
FEATURES		Seq primer: -40UP from G1bco					
		High quality sequence stop: 409.					
		Location/Qualifiers					
		1..454					
		/organism="Homo sapiens"					
		/note="Organ: pancreas; Vector: PCMV-SpOrf6, site_1: SalI;					
		site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.					
		* Average insert size 1.72 kb. Life Technologies catalog #:					
		11548-013"					
		/db_xref="taxon:9606"					
		/map="5q35.3"					
		/clone="IMAGE:2223057"					
		/clone_1id="NCI-CGAP_Pan1"					
		/tissue="type="adenocarcinoma"					
		/lab_host="DH10B"					
BASE COUNT		95 a	121 c	127 g	110 t	1	others
ORIGIN							
		Query Match	8.8%;	Score 215;	DB 28;	Length 454;	
		Best Local Similarity	86.9%;	Pred. No. 1.06e-257;			
		Matches 246;	Conservative	0;	Mismatches 37;	Indels 0;	Gaps 0;
Db	8	GAGTTTGCTCTGTGTGCCAGCGCTGGAGTGCAGATGCGCATCTCGGCTCATTTGCAACC	67				
OY	2420	gagttgtcctctgttaccacagcgcggaggtgaatagcgcgcttctgggctaccgcgaacc	2361				

Db	363	TAGAGACGGGGTTCCTCCAGTGTGGTCAGGCTGCTTCGAATCCCGACCTCAGGTATC	422
Cp	2262	tagagatgggggtttctccacgttgcacgttcgacgttcgttgaactcccgcccttaagatgac	2203
Db	423	CGCCACCTCGGCTGCCAAGAGTGTAGAGTTACAGGATGAGCCACCGCCGCCGCCCA	482
Cp	2202	caccacacttgcctcccaagaagtgttgagattacagtgcttgagccacgtcgccctggccaa	2143
Db	483	GA 484	
Cp	2142	ga 2141	
RESULT	14		
LOCUS	AC098015	347 bp	DNA GSS 26-AUG-1998
DEFINITION	HS.3041.A1.H02.MF CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone plate-3041 Col-3 Row-O, genomic survey sequence.		
ACCESSION	AC098015		
NID	G3469044		
VERSION	AC098015.1	GI:3469044	
KEYWORDS	GSS.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryote; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.		
AUTHORS	1 (bases 1 to 347)		
TITLE	Mahairas,G.G., Zackrone,K.D., Smith,T., Tipton,S., Schmidt,S., Tralcoff,R., Adjajan,C., Blanchard,A., West,A. and Hood,L.E. Construction of a Characterized Clone Resource for Genomic Sequencing: Generation and Preliminary Analysis of 20,000 Sequence Tagged Connectors		
JOURNAL COMMENT	Unpublished (1997)		
FEATURES	Source		
FEATURES	High quality sequence stop: 347.		
FEATURES	Location/Qualifiers		
FEATURES	1..347		
FEATURES	/organism="Homo sapiens"		
FEATURES	/note="Organ: Sperm; Vector: pBelOBAC11; BAC Clones in E-Coli DH10B"		
FEATURES	/db_xref="taxon:9606"		
FEATURES	/clone="plate-3041 Col-3 Row-O"		
FEATURES	/clone_lib="CIT Approved Human Genomic Sperm Library D"		
FEATURES	/sex="male"		
BASE COUNT	62 a 103 c 80 g 102 t		
ORIGIN			
Query Match	8.7%: Score 211; DB 39; Length 347;		
Best Local Similarity	85.0%: Pred. No. 8,29e-252;		
Matches	249; Conservative 0; Mismatches 44; Indels 0; Gaps 0;		
Db	54	TTTTTTTTTTTTTTTGAGATGAGATTTCGCTCTTCGTGCTCAGGCTGAGTGAATGCG 113	
Cp	2442	TTTTTTTTTTTTTTTTTTTGGAGAGAGTGTGCTCTTGTCTGTACCCAGCTGAGTGAATGCG 2383	
Db	114	GCAATTCGGCTCACCACCACTTCACCTCCAGGTTCAAGCAATTTCTCTTGCCCTCAGGC 173	
Cp	2382	gcgntctggctcaccacgaacctctgcctcccaagttcaagaaatnttctgcctgcgcg 2323	
Db	174	TCACGAGTAGCTGGGATTACAGAGCATGGGCGACCATGCGCGGCTAATTTGATTTGAG 233	
Cp	2322	tccctgagtagcagaagttacagcatgctgcctaccacgcccggcgaatttngatttttng 2263	

Db	234	TAGAGAGGGGCTTCCACGATGTGGTCACGCGGTGCTCCAACTCCAGCTCAGGATGC	293
Cp	2262	tagagaggggttcttcacagttggtccagagcgggtcttgaaectccggcctcatgatacc	2203
Db	294	CACCGCGCTCGGCCCCCAAGTGTGGATTACAGGTGTGAGCACCTCGCC	346
Cp	2202	caccaccttgacctcccaagaagtgtggattacaggtgtgagccactgcgc	2150
RESULT	15	AQ167755	397 bp DNA GSS 15-OCT-1998
LOCUS		HS-2198_A2.G10.MF.CIT	Approved Human Genomic Sperm Library D Homo
DEFINITION		sequence genomic clone Plate-2198 Col-20 Row-M, genomic survey	
ACCESSION		AQ167755	
NTD		93561470	
VERSION		AQ167755.1	GI:3561470
KEYWORDS		GSS.	
SOURCE		human.	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;	
AUTHORS		Primates; Catarrhini; Hominiidae; Homo.	
TITLE		1 (bases 1 to 397)	
JOURNAL		Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Furlong,J.,	
COMMENT		Shaker,R., Schmidt,S., Traicoff,R. and Hood,L.E.	
		Construction of a Characterized Clone Resource for Genomic	
		Sequencing	
		Unpublished (1998)	
		Contact: Mahairas GG, Wallace JC, Hood L	
		High Throughput Sequencing Center	
		University of Washington	
		401 Queen Anne Avenue North, Seattle, WA 98109, USA	
		Tel: (206) 616-3618	
		Fax: (206) 616-3887	
		Email: jwallaceu.washington.edu	
		Sequence Tagged Connector	
		Plate: 2198 Row: M Column: 20	
		Class: BAC ends	
		High quality sequence stop: 397.	
FEATURES		Location/Qualifiers	
source		1..397	
		/organism="Homo sapiens"	
		/note="Organ: sperm; Vector: pBelOBAC1; BAC Clones in	
		E-Coli DH10B"	
		/db_xref="taxon:9606"	
		/clone="Plate-2198 Col-20 Row-M"	
		/clone_1fb="CIT Approved Human Genomic Sperm Library D"	
		/sex="male"	
BASE COUNT		106 a 98 c 116 g 76 t 1 others	
ORIGIN			
Query Match		8.7%; Score 212; DB 40; Length 397;	
Best Local Similarity		84.5%; Pred. No. 2,79e-253;	
Matches	251; Conservative	0; Mismatches 46; Indels 0; Gaps 0	
Db	52	GGCGAGGTGTGGTCTCAGCGCTGTATCCACACTTGGAGCGTGAAGTGGTGA	111
Qy	2145	ggccagggcgcaatgctcacacctgtaatcccaaccttggagggccaaggtgggga	2204
Db	112	TCACCTGAGGTGGAGTTCAGACACAGCTGACCACATGAGGAACCCCGCTCTACT	171
Qy	2205	tcattcaagccggaggtcaagaaccagcctgaccaacgltggaaacccatcttact	2264
Db	172	AAAAATACAAATTTAGCCGGACATGTGTGGCCATGCGCTGTATCCAGCTACTCTGGGAGG	231
Qy	2265	naaaatacnaaatagccgggagctggtagcgcaatgctgaanctggtctcacaagaag	2324
Db	232	CTGAGCGAGGAATACGCTTGTGAAACCCGGAGCGGAGCTCCGATGACTGATCGGC	291
Qy	2325	ccgagggcgaanaatttcttgacatcggggagcgagggttcggtgagcccgannggc	2384
Db	292	CATTGCATCCAGCTGNGCAACAAGACCAAACTCCTCTCAATTAAGAAAAAAGAA	348

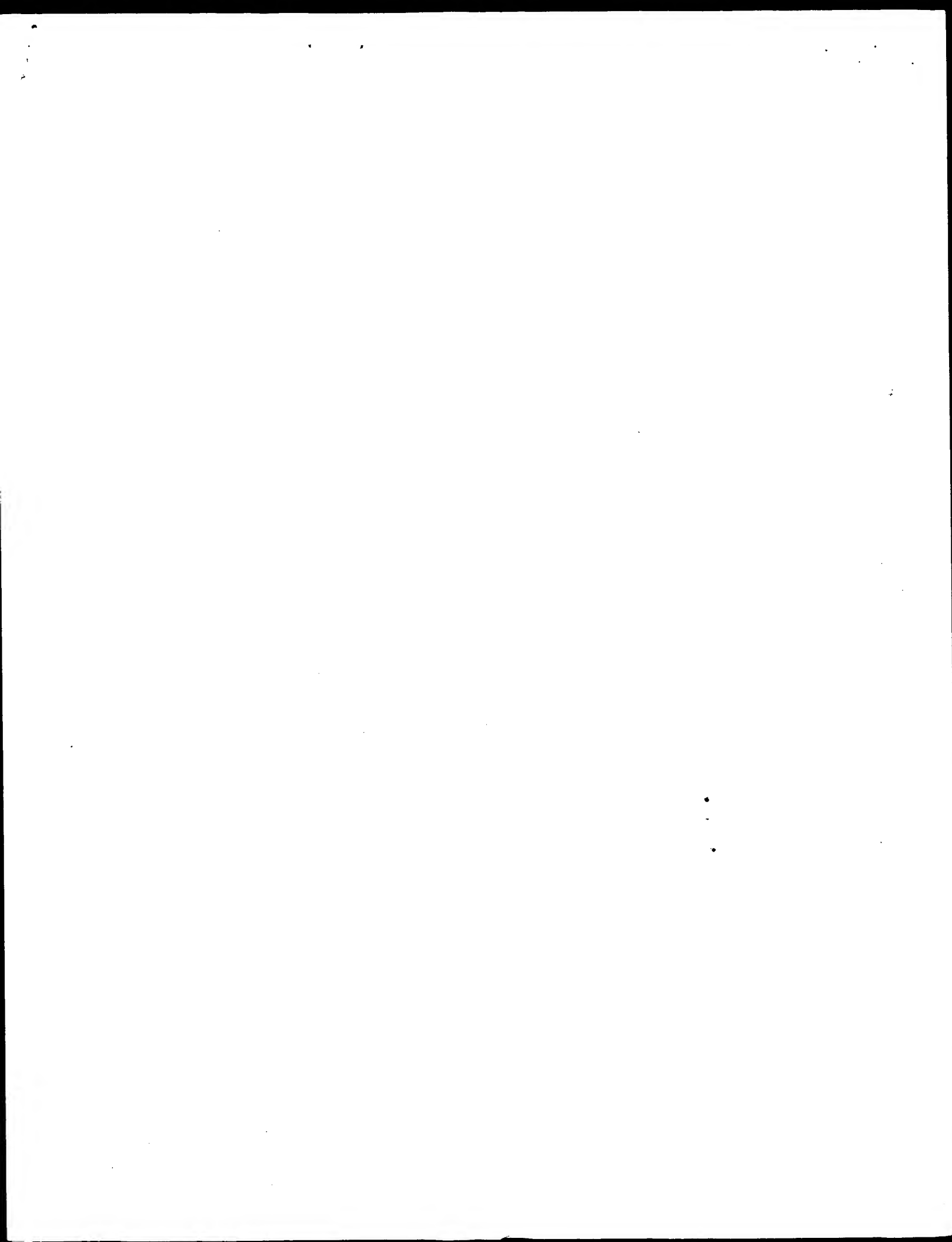
Thu Jan 20 11:03:51 2000

US-09-246-129-1.rst

Page 9

|||||
QY 2385 catgcactccagcctgggtacacaagagcaaacctctgtccaaaaaa 2441

Search completed: Tue Jan 18 15:44:07 2000
Job time : 4186 secs.



FT /note= "base n at position 2341 is not identified
 in the specification"
 FT misc-difference 2379
 FT /*tag= 1
 FT /note= "base n at position 2379 is not identified
 in the specification"
 PN WO9614328-A1.
 PD 17-MAY-1996.
 PF 07-NOV-1994; U12880.
 PR 21-NOV-1994; WO-012880.
 PR (HMAE-) HUMAN GENOME SCI INC.
 PI N1 J, Rosen CA, Yu G;
 DR WPI; 96-321550/32.
 DR P-RSD: R99453.
 PR Tumour necrosis factor gamma and corresponding DNA and RNA - for
 PT treating cancer, facilitating wound healing and providing resistance
 PT against infection
 PS A claim 1; Fig 1; 69p; English.
 CC A CDNA clone (T3030), deposited as ATCC 75927, codes for tumour
 CC necrosis factor-gamma (R99453), a novel member of the TNF family. It
 CC was isolated from a human umbilical vein endothelial cell CDNA
 CC library. The coding sequence can be incorporated into a vector and
 CC used for prodn. of recombinant TNF-gamma using bacterial, insect or
 CC mammalian cells as hosts. It may also be used as a diagnostic probe,
 CC for gene mapping, or for the gene therapy of diseases involving
 CC abnormal cellular proliferation, e.g. tumours and cancer. Antisense
 CC sequences are useful as TNF-gamma antagonists.
 SO Sequence 2442 BP; 760 A; 529 C; 511 G; 636 T;
 Query Match 99.8%; Score 2430; DB 21; Length 2442;
 Best Local Similarity 99.9%; Pred. No. 0.00e+00;
 Matches 2439; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Db 601 ttctcaagtcagaccactcagctcctcaagaagcgaagtattgcccaggtcactagtc 660
 QY 601 ttctcaagtcagaccactcagctcctcaagaagcgaagtattgcccaggtcactagtc 660
 Db 661 caagaatgtatctcttgaacaaatgtgtatgtccagtcacatatcttcattctcc 720
 QY 661 caagaatgtatctcttgaacaaatgtgtatgtccagtcacatatcttcattctcc 720
 Db 721 tcccaagaagctttttagcgttagatattcagatcatttagctatttgaatgt 780
 QY 721 tcccaagaagctttttagcgttagatattcagatcatttagctatttgaatgt 780
 Db 781 atatagaagctttttagcgaagcttcaagcttcccaatgagaatattatctcttc 840
 QY 781 atatagaagctttttagcgaagcttcaagcttcccaatgagaatattatctcttc 840
 Db 841 ttgtcttccagctgtgagacaaactcccaacagcatttaaaatcagttcccaagtc 900
 QY 841 ttgtcttccagctgtgagacaaactcccaacagcatttaaaatcagttcccaagtc 900
 Db 901 tgcactgggaactagaactagcctggtcctcacaagaaccgaatgacatataccaca 960
 QY 901 tgcactgggaactagaactagcctggtcctcacaagaaccgaatgacatataccaca 960
 Db 961 aattcctgtgattcccaagctcggagagacttacttacttcccaagtcacattccgtg 1020
 QY 961 aattcctgtgattcccaagctcggagagacttacttacttcccaagtcacattccgtg 1020
 Db 1021 ggaatgactctgagctgagctggaatcaacagaagcggcggcgaagaagcgaatcca 1080
 QY 1021 ggaatgactctgagctgagctggaatcaacagaagcggcggcgaagaagcgaatcca 1080
 Db 1081 tcaactgtgtatcccaagctgaagcagcagctacccctgaagcgaacccaagcttcctatg 1140
 QY 1081 tcaactgtgtatcccaagctgaagcagcagctacccctgaagcgaacccaagcttcctatg 1140
 Db 1141 ggaacaaagctgtatgagaaagctgaagctgaagcttcccaagcgaacccaagcttcctatg 1200
 QY 1141 ggaacaaagctgtatgagaaagctgaagctgaagcttcccaagcgaacccaagcttcctatg 1200
 Db 1201 ttttctccttgaagaagcggagcgaagctgaatgtgaaagctgaagcgaagcgaagcgaagcga 1260
 QY 1201 ttttctccttgaagaagcggagcgaagctgaatgtgaaagctgaagcgaagcgaagcgaagcga 1260
 Db 1261 attacacaaagaagataaaactcttcttgagaccccttctactatagagagagcaaat 1320
 QY 1261 attacacaaagaagataaaactcttcttgagaccccttctactatagagagagcaaat 1320
 Db 1321 atcaatatagaatcctcctgacacagcttcccaatttcttctgttccaatgttaata 1380
 QY 1321 atcaatatagaatcctcctgacacagcttcccaatttcttctgttccaatgttaata 1380
 Db 1381 taacacaggggttttcttcttgagcggagagagcattcccaagcgaagcgaagcgaagcgaagcga 1440
 QY 1381 taacacaggggttttcttcttgagcggagagagcattcccaagcgaagcgaagcgaagcgaagcga 1440
 Db 1441 tatgaatcttgaggccaataattcaccacttcagtgccttactgatagagactactagtc 1500
 QY 1441 tatgaatcttgaggccaataattcaccacttcagtgccttactgatagagactactagtc 1500
 Db 1501 gaaaagcgtggaagagagaataatataatattgaatgtgttgaagattgaggatttct 1560
 QY 1501 gaaaagcgtggaagagagaataatataatattgaatgtgttgaagattgaggatttct 1560
 Db 1561 aaatattaagacactgtatcaataaagatgagatgagatcaccacgggttcaggattgaaga 1620
 QY 1561 aaatattaagacactgtatcaataaagatgagatgagatcaccacgggttcaggattgaaga 1620
 Db 1621 gaaataattcaaacactcctgctatbacaatgtgtcaacagctgtgtccagttattgttcaat 1680
 QY 1621 gaaataattcaaacactcctgctatbacaatgtgtcaacagctgtgtccagttattgttcaat 1680

DE	05-JUN-1998	(first entry)
DE	Genomic DNA for interferon-gamma production inducer.	
KW	Interferon-gamma; IFN-gamma; production inducer; gene therapy;	
KW	Immunocompetent cell; treatment; prevention; malignant tumour;	
KW	Viral infection; bacterial infection; immune disease; ds.	
OS	Homo sapiens.	
FH	Key	Location/Qualifiers
FT	CDS	/tag= a 15607..27049
FT	sig_peptide	/tag= b 15607..20468
FT	mat_peptide	/tag= c 20469..27046
FT	intron	/tag= d 15686..17056
FT	intron	/number= 1 17069..20451
FT	intron	/tag= e /number= 2 20587..21920
FT	intron	/tag= f /number= 3 22055..26827
FT	intron	/tag= g /number= 4 /tag= 9
FN	EP-816499-A2.	
PD	07-JAN-1998.	
PF	27-JUN-1997:	304616.
PR	27-JUN-1986:	JP-185305.
PL	(HAYB) HAYASHIBARA SEIBUTSU KAGAKU.	
PA	Kurimoto M, Okura T, Torigoe K;	
DR	P-PSTDB: W477429.	
PT	Genomic DNA encoding polypeptide inducing interferon-gamma	
PT	production - by immuno-competent cells, useful to treat e.g. human	
PT	malignant tumours or viral diseases	
PS	Claim 13; Pages 64-73; 74pp; English.	
CC	The present sequence encodes a protein, which induces	
CC	interferon-gamma (IFN-gamma) production in immunocompetent cells.	
CC	The protein has high biological activity, including enhancing	
CC	cytotoxicity of killer cells and inducing killer cell formation,	
CC	in addition to inducing IFN-gamma production by immunocompetent	
CC	cells when expressed in mammalian cells, facilitating its use in	
CC	low dosages to treat/prevent, e.g. malignant tumours, viral or	
CC	bacterial infections and immune diseases. As it is expressed in	
CC	mammalian cells, it also has low toxicity when used in human	
CC	treatments, minimising side effects. The DNA encoding the protein	
CC	can be used in gene therapy, e.g. by injecting vectors containing	
CC	the DNA or transplanting cells.	
SQ	Sequence	28994 BP; 9141 A; 5655 C; 5654 G; 8544 T;
Query Match	9.2%; Score 223; DB 39; Length 28994;	
Best Local Similarity	86.3%; Pred. No. 9,84e-124;	
Matches	258; Conservative	0; Mismatches 41; Indels 0; Gaps 0
Db	23428	tggcggcggtgctggccaacacctgaatccagcactttgggaggcggaggggcgg 23487
Qy	2144	tggccaagcgcgctggcccacaccctgatatacccaacttggggagcgcaaggctgggtgg 2203
Db	23488	atcaccttgaaatcagagattcaagaacagccctgcaccaatgagaaaaccatctaac 23547
Qy	2204	atcatctaagcggcgagtccaagacagccctgcaccaagctgagaaaaccatctaac 2263
Db	23548	taaaaataaaaaaatagctcggtcggtggcgataagctgttaattccccgactatctggag 23607
Qy	2264	taaaaataaaaaaatlagccggcggtggtatgccatgctgttaanccttcggctactaaggag 2323
Db	23608	gctggcgcgaggaatctttaaccgggagacagagtttcagatgagcctlagaatcgtg 23667
Qy	2324	gccggcgagcgaanaattincttgaacttgggagagcgaggttgcgttggccagangcgg 2383
Db	23668	cattgcactcagccttgaggcaacaagacaaactcgtgtcaaaaaaaaaaaaaaaa 23726

QY	2384	cattgacgtcccaaccctggtaaacagaagcacaactctgtccaaaaaaataaaa	2442
RESULT	4		
ID	092779	standard; DNA; 6905 BP.	
AC	092779;		
DE	09-DEC-1995	(first entry)	
DE	Human thymopoleitin continuous gene fragment.		
KW	Thymopoleitin; diagnostic; therapeutic; gene therapy; ss.		
OS	Homo sapiens.		
FT	Key	Location/Qualifiers	
FT	exon	39..1956-2843	
FT		/tag= a	
FT		/note= "exon 4"	
FT	exon	4691..4788	
FT		/tag= b	
FT		/note= "exon 5"	
PD	WO9517205-A1.		
PN	29-JUN-1995.		
PR	13-DEC-1994;	U14356.	
PR	21-DEC-1993;	US-171382.	
PA	(IMMU-) IMMUNOBIOLOGY RES INST INC.		
PI	Culler MD, Goldstein G, Harris CA, Setcavage DR;		
PI	Shenbgamurthi P, Steklerka JJ, Talle MA,		
DR	WPI: 95-240474/31.		
PT	Polynucleotide(s) encoding human thymopoleitin proteins - used to		
PT	develop prod. for diagnosis and therapy involving immune or nervous		
PT	system conditions.		
PS	Disclosure: Fig 9a-9g; 85pp; English.		
CC	The sequence represents a fragment of the human thymopoleitin gene,		
CC	including exons 4 and 5. DNA encoding human thymopoleitin, from		
CC	a cDNA library prepared from human thymus RNA, may be expressed in		
CC	a host cell e.g. bacterium (preferably Escherichia coli), fungus,		
CC	insect or mammalian cell for production of recombinant thymopoleitin.		
CC	Thymopoleitin has a regulatory effect on the mammalian immune system,		
CC	and can be used for treating/modulating an immune or nervous system		
CC	condition, and for the treatment of chronic infection, autoimmune		
CC	disorders and certain affective psychiatric or neurological		
CC	disorders. The protein and the DNA can be used as diagnostics;		
CC	the DNA can be used in gene therapy.		
SQ	Sequence 6905 BP; 2028 A; 1280 C; 1359 G; 2238 T;		
Query Match	8.8%; Score 214; DB 15; Length 6905;		
Best Local Similarity	84.9%; Pred. No. 9,21e-118;		
Matches	255; Conservative 0; Mismatches 45; Indels 0; Gaps 0;		
Db	6171 ggcggcgccagatggtcgaagcgttgtaatcccagcactttggaggccgggccccgga	6230	
QY	2145 gccggcgccagtgtgctcacaccttgaatcccaactttggaggccaaggttgggga	2204	
Db	6231 tcaactgggttcgcgaagttcgaagccgagccgcgcacacatggaagaaaccctgtctact	6290	
QY	2205 tcattaaagccgggagttcaagaccagcctgaccacaagtggaagaaaccatcttact	2264	
Db	6291 aaaatacaaaattagccagcatgtgtgtgtcatgtcctgtaatccctagctactccgaag	6350	
QY	2265 naaaatacaaaattagccgggcgtgtgtagcgcgatgctgtaanctgtctactcaagga	2324	
Db	6351 ctgaagcagggagaatcgccttgaaccttggaggccggaggttgtctgtgcccgaagtcgcgc	6410	
QY	2325 ccgagccaggaanaattctnctgaacttgaagtggaagccagaggttgcgtgagccgaagccgc	2384	
Db	6411 cattgcaacttgaagtcgggacaacagaagcgaactctgtccaaaaataataataa	6468	
QY	2385 cattgcaactccagcctgtgttaacaagagcaaatctgttccaaaaaaaaaaaaa	2442	
RESULT	5		
ID	V02032	standard; DNA; 40352 BP.	
AC	V02032;		
DT	12-JUN-1998	(first entry)	
DE	MAGE-B cluster DNA sequence.		
KW	MAGE-B cluster DNA; testicular seminoma; cancer; screening; leukemia;		

[illegible][illegible]

/db.xref="PID:94038081"
/db.xref="GI:4038081"
/translation="MRPRLSKVYSEPMKRLILFVPEVQVOTPHQKNOPLAHMEH
EDGLAFTNNKNTYNNKFLIPESGDYFIYSQVTRGMTSECELRQGRNKPDSITV
VITVTDSEPEPTOLMLGTKSVCEVSNWEPQIPLGAFSLQESQDKLMVSDISLSD
YTKEDKTFEFGAFLL"
BASE COUNT 854 a 603 c 577 g 750 t 1 others
ORIGIN

Query Match 98.2%; Score 2393; DB 31; Length 2785;
Best Local Similarity 99.1%; Pred. No. 0.00e+00;
Matches 2423; Conservative 0; Mismatches 18; Indels 3; Gaps 3;

Db 340 CCAAAATCAGAGAAATTCATCTACTATTCACAGTTGGCCAACTTCCAGTCTAGAGACA 399
QY 1 cccaatcaagaagaatccatcatcactacacagcttgcgcagcttcccaagctcagtgccaga 60
Db 400 AATCCAGGACCTCAGACCTAGAGTTCTTATCTGAGACTCCAGAGAAAGAACAA 459
QY 61 aatccaaggcacctcacactcagagcttccctacactctcagctccagagagaagaacaa 120
Db 460 GACAGTGCAGAAAGATATTTAGAACCCCTGAAACCTAGAGTTTAAAGAGAGAT 519
QY 121 gacagtgacagaagatagtagaacccacgaaacctagaaggtgaaagagagat 180
Db 520 ACCCTCTGACCTATPAGAAATTTTCAAGTTCAGAGGGGATATCTTGTGGCCCAAGAC 579
QY 181 accctctgactataagaatcttcagctgcaggggagatatactctgtgcccagaac 240
Db 580 ATTGCTGTTATCATTTGACTTAAGAGAAATTTTGTGCTGAGCTCTGAGTGAAGATTAG 639
QY 241 attgctgtatactattgactaagaagaatattctgtgtgagccctgagtagattag 300
Db 640 GACAGGAGAGAGGAGGAGTTTATCTACTTACTGCTGCTAGAGAGAGTTTGT 699
QY 301 gaccagggagagagcagagcttccatcactcactcagctcgtgaagagaggttct 360
Db 700 CCAATGATGAATGGGAGTAAAGATTTGAGCAATCACTTATGGGCAAGAGGAGAA 759
QY 361 ccaattgagatgggagagagaaacagctcagccaactcactatggg-cacagaaagg-aa 418
Db 760 TTTGAAGGCTTGTGCTGCTGCTGCTGCTACGTAAGAGAGAGCATGAGATT 819
QY 419 ttggaaggctgtgtgtcgcctcgttcatacagtaacagaagagagctcgtgagttt 478
Db 820 ATCTGATGATTTGGGAGAGATTTCTTGAGAGAGCATTTCTTAAACAGAGAGA 879
QY 479 atctgagatcttgggagagagaaatcttcgacacagccattctcctaacaacagaga 538
Db 880 ATAGGGGATTCCTTAACCTTCATTTCTCCAGATCATAGGTCAGAGTAAATTTAA 939
QY 539 ataggggattccttaactcattcgttctccagagatcaggtcctcagatataat 598
Db 940 AATTTTCAAGTCAAGACCTCAGTCTCAGAAAGGCAAGTAATTTTCCCAAGTCACTAG 999
QY 599 aatttcaagtcagacactcagctcagaaagaaagtaatttggcccaggtcactag 658
Db 1000 TCCAAGATGTTATCTTCTTGAAGAAAGTATGTCGACATATCTTCAATTCATT 1059
QY 659 tccaagatgattcttcttgaacaaagtcagtcagccagtcacatcttccatcattc 718
Db 1060 CCTCCCAAGAGAGTTTATAGTGTAGTATATTCGATCACTTTAGTTTGAAGAA 1119
QY 719 cctcccaagaagctttagttagctgtagtatattcagcactttagtctatttgagaa 778
Db 1120 TGATATGAGAGCTTTTAAAGCAAGTCTACAGTTTCCATGAGAGAAATTAATCTCTT 1179
QY 779 tgatagagagcgtttttaaagcaagctacaggttcccaatgagagaaatlaactcctt 838
Db 1180 TCTTGTCTTTCAGTGTAGAGCAAACTCCACAGCACTTTAAATAATCACTTCCAGC 1239
QY 839 tctgtcttccagctgtgagacaactcccaacagcactttaaataacagttccccagc 898

Db 1240 TCTGACATGGGAGACATGAGTACGAGCTGGCTGCTTACAGAAAGACGAGTATACCA 1299
QY 899 tctgacatgggagacatgagactagagccttgcttccacaaagacagatagactatac 958
Db 1300 CAAATTCCTGCTGATCCAGAGTGGGAGACTTCTTATTTACTCCAGGTACATTCG 1359
QY 959 caaatctcgtgatcccgagctcgagagactactcattcactcccgatcactccg 1018
Db 1360 TGGGATGACCTCGATGAGAGTGAATTCAGACAGAGGCGACCAACCAACCGAGACT 1419
QY 1019 tgggagacccctcgagtgagtgaaatcagacagagcgacacaaagccagactc 1078
Db 1420 CATCACTGTGTCATACCAAGGTAACAGACAGCTTACCTGAGCCAAACCACTCTCAT 1479
QY 1079 catcactgtgtatataccaaggttaacagacagctacccctggccaaccagctctcat 1138
Db 1480 GGGGACCAAGTCTGTATGCGAAGTACGTAGCAACTGTGTCAGCCATCTACTGGAGC 1539
QY 1139 ggggaccaaagctgtatgagagtagtagcaactgttccagccatctactcctgagc 1198
Db 1540 CATGTTCTCTTGCAGAGAGGAGCAAGCTAATGTGAGACGTACAGTATCTTGTGT 1599
QY 1199 catgtctctctcgagagagggagagactaatgtgagcgcagtgacatctcttgt 1258
Db 1600 GGAATTACAAAAGAGATTAACCTTCTTGGAGCCTTCTTACTATAGAGAGAGCA 1659
QY 1259 ggaattacaaaagagatataaaccttcttgagccttctactatagagagagcaa 1318
Db 1660 ATATCATTAATGAAGTCTCTGCGACCGAGTCTCTAATTTCTTGTTCAAATGTA 1719
QY 1319 atatcatatagaagctctctgcacacagagcttcttaattcttctgtcctaagta 1378
Db 1720 TATACAGAGGGTCTTCTTGGGCGGAGTAGGGGAGTCCACAGAGAGCAAGGTTTA 1779
QY 1379 tatacagaggttctctctgtggcgagtagggagcttccacagagagcaaggttca 1438
Db 1780 GCTATGAATTTGGGCGCAAAATTTTACACTTCACTGCTTACTGATGAGTAAAC 1839
QY 1439 gctatgaatttggggccaaaatttcacactcattgtgcttactatgaggtactaac 1498
Db 1840 TGGAAAAGGCTGAAGAGCAATATATTAATGAGTGGAGAGTGGCGAGTT 1899
QY 1499 tggaaaaggctgaagagagcaaatatataagaaggttggagagcttgcgagttc 1558
Db 1900 CTAAATTTAAGACACTGATCAATGAATGAATGATGATCTCGGCTCAGAGATTGA 1959
QY 1559 cttaatttagacactgactactaataatgagatgagactactcgggtcagattgaa 1618
Db 1960 GAGAAATATTCAACACCTCC-TGCTATACATGTCACCAAGTGTCAAGTTATGTTCA 2018
QY 1619 gagaatatttcaacacctccctgcatacaatgttcacacagtggtccagttatgtca 1678
Db 2019 ATTGATCATTAATTTGCTTCAATTCAGAGGCTTGAAGAGTCCAGAAAGCTCTAG 2078
QY 1679 attgacataaatttgcctcaatcagagccttgaagaaagtcacaaagagcctag 1738
Db 2079 AAAACGATTAACCTTTCAGAGCAAAATCTTCAACAAATTTTCCACATTTCTCATGC 2138
QY 1739 aaaaagataaacttccagaggaagaatcttcacacatttccacacttcttaagc 1798
Db 2139 CTTCGCTAAAAAATGAAGAGAGTGTGTATGTCTCATGAATGTTCAACAGAAAGAG 2198
QY 1799 ctctccttaaaaaatgaagagaggtgtatgtctctcagatgttccacagaaagag 1858
Db 2199 TTGGTTTTCATGTCATCTACAGATATGAGAAAGTACCTTTCTTATTAATGTAC 2258
QY 1859 ttggttttcatgcatctacagacatagaaagctactcttcttgaatagtaac 1918
Db 2259 AGATATTAATGAAGAGATGATGATTTTCACTATATCAAAATTAACAAAGTTGCTG 2318
QY 1919 agatatctaataagagaggttgaagttcacaatgtatatacccaataacaggttgc 1978
Db 2319 TATTCAGTAGAGTTTCTTTCGCCACCTAATTTGTGCTGGGTTCTACCTTAACCCAGAGA 2378


```

repeat_region .26986 .26956
/note="MIR repeat: matches 145. .84 of consensus"
repeat_region 29028. .29139
/note="MIR2 repeat: matches 31. .143 of consensus"
repeat_region 29280. .29388
/note="MIR repeat: matches 13. .120 of consensus"
repeat_region 29374. .29473
/note="MIR repeat: matches 158. .259 of consensus"
repeat_region 29821. .29920
/note="MIR repeat: matches 261. .163 of consensus"
repeat_region 29913. .30022
/note="MIR repeat: matches 135. .9 of consensus"
repeat_region 30034. .30352
/note="L1H5 repeat: matches 247. .571 of consensus"
repeat_region 30155. .30352
/note="L1P3 repeat: matches 359. .571 of consensus"
repeat_region 30741. .31016
/note="AluYo repeat: matches 295. .1 of consensus"
repeat_region 31461. .31645
/note="MIR repeat: matches 70. .261 of consensus"
repeat_region 31732. .31863
/note="MIR repeat: matches 18. .154 of consensus"
repeat_region 31879. .32097
/note="L1MBA repeat: matches 904. .672 of consensus"
repeat_region 33492. .33680
/note="MIR repeat: matches 240. .30 of consensus"
repeat_region 33756. .33895
/note="MIR2 repeat: matches 145. .1 of consensus"

```

Note: remainder of annotations omitted.

Query Match	9.5%;	Score 231;	DB 29;	Length 123004;
Best Local Similarity	87.4%;	Pred. No. 3.30e-153;		
Matches	263;	Conservative	0;	Mismatches 38;
			Indels	0;
			Gaps	0

Db	49966	TTTTTTTTTTTTTTTTTTGAGACGGAGTTTCGCTCTTTTGGCCAGCGGTGGAGTGCATAAGC	50025
Cp	2442	TTTTTTTTTTTTTTTTTTTGGACAGAGATTTCCTCTTGTACCCAGCTGTGAGTGTCAATGTC	2383
Db	50026	GGCATCTGGGCTCACCGCAACCTGCGCTCCCGGGGTTCAGACGATTCCTCTCTCACGCC	50085
Cp	2382	GGGATCTGGGCTCACCGCAACCTGCGCTCCCGGGGTTCAGACGATTCCTCTCTCACGCC	2323
Db	50086	TCCCGATGAGCTGGGATTACAGGGCATGGCGCCACGACGCCACGAGTAAATTTTGTATTTTAA	50145
Cp	2322	TCTGTGTAGTCAAGTTAAGCAATGTGCTGCTACCAAGCCGCGTGAATTGTTATTTAG	2263
Db	50146	TAGAGACGGGGTTTCTCCACGTTGGTCAAGCTGTCTCGAATCTCCACACTAGCGCATC	50205
Cp	2262	TAGAATGTGGTGTCTTCCACGTTGGTCAAGCTGTCTCGAATCTCCACACTAGCGCATC	2203
Db	50206	CGCCCGGCTTTGGCTTCCCAAAGTGTGGGATTATAGAGATGAGACGACACTCGGCCAGGCCA	50265
Cp	2202	CACCCGCTGTGGCTTCCCAAAGTGTGGGATTATAGAGATGAGACGACACTCGGCCAGGCCA	2143
Db	50266	G	50266
Cp	2142	G	2142

RESULT	4	AC006286	170820 bp	DNA	HTG	30-MAR-1999
LOCUS		Homo sapiens, clone hRPK.21.A.1, WORKING DRAFT SEQUENCE, 2 ordered				
DEFINITION		pieces.				
ACCESSION		AC006286				
NID		94544485				
VERSION		AC006286.9	GI:45444485			
KEYWORDS		HTG; HTGS_PHASE2.				
SOURCE		human.				
ORGANISM		Homo sapiens				
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;				
		Eutheria; Primates; Catarrhini; Homiidae; Homo.				
REFERENCE		1 (bases 1 to 170820)				

AUTHORS
TITLE
Homo sapiens, clone hRPK_21_A_1
JOURNAL
REFERENCE
AUTHORS
2 (bases 1 to 170820)
Britten, B., Linton, J., Nussbaum, C., Lander, E., Allen, N., Anderson, M.,

Baker, J., Baldwin, J., Barna, N., Beckett, R., Benn, J., Brown, A.,
 Castle, A., Cerny, J., Colangelo, M., Collins, S., Collymore, A.,
 Cooke, P., Dearellano, K., Depyre, E., Devon, K., Dewar, K.,
 Donnell, L., Ferreira, P., FitzHugh, F., Forrest, C., Funke, R.,
 Gagne, D., Galagan, J., Gardyna, S., Grant, G., Hagos, B., Harford, A.,
 Herena, L., Horton, L., Howland, J. C., Jones, C., Kann, L., Karatas, A.,
 Lehoccky, L., Macdonald, P., Margolis, N., Mcwan, P., McGurt, A.,
 McKernan, K., Meldrum, J., Molla, M., Morris, M., Morrow, J.,
 Mychalekcy, J., Naylor, J., Nlloff, M., O'Connor, T., O'Donnell, P.,
 Pavlin, B., Peterson, K., Pollara, V., Riley, R., Roberts, D., Roy, A.,
 Severy, P., Stange-Thomann, N., Stojanovic, N., Stone, C.,
 Sudbramanian, A., Tesfaye, S., Torruella-Miller, I., Vassiliev, H.,
 Vo, A., Wagner, A., Wheeler, J., Wu, X., Wyman, D., Ye, W.-J. and Zody, M.
 Direct Submissions
 Submitted (05-JAN-1999) Whitehead Institute/MIT Center for Genome
 Research 320 Charles Street, Cambridge, MA 02141, USA
 On Mar30, 1999 this sequence version replaced gi:4454493.

http://ftp.genome.washington.edu/RM/RepeatMasker.html.

```

* consists of 2 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor. This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved
*
*       1  96189: contig of 96189 bp in length
*           gap of unknown length
*       96190 170820: contig of 74631 bp in length.

```

FEATURES	Location/Qualifiers
source	1. .170820

BASE COUNT	ORIGIN
50750 a	36911 c 37027 g 46132 t

Query Match	9.5%	Score 232;	DB 20;	length 170820;
Best Local Similarity	87.7%;	Pred. No. 5.14e-154;		
Matches 265;	Conservative	0;	Mismatches 37;	Indels 0;
				Gaps 0;

Db 148792 TTGGCCGGGGCCAGTGGCTACCCCTGTAAATCCAGACTTTGGGAGGCCAAGCGAGCA 148851
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
Oy 2143 ttggccaggcgcaatgctacacacctgtaatcccaacaccttgggagggccaagtgggtg 2202

Dd 148952 GATCACTTGAGACACMGCGAGTTCGAAGACCAGCCTGACTACATGTGAGAAGCCCCCATCTCTA 148911
||||| - - - ||||| - - - ||||| - - - ||||| - - - ||||| - - - ||||| - - - ||||| - - - |||||
Oy 2203 gatcatctaagggccgggagtctcaagaaccagccttaaccaactgtgagaaccccccatctcta 2262

Db 148912 CTAATAATACAAATTACTCTGGGCTTGCGCACATGCCTCTAATCCAGCTACTCAGA 148971
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Oy 2263 ctnaaaatacmaaatlttagccggcgctgttgcgcatgcgtctaancctgctactaccaga 2322

Db 148972 GCGTCAGGCAGAGANTCGCTTGAACCTGGGACGACAGAGTTGGCGGTAGCCGACATGCC 149031
 ||| ||||| ||| ||| ||||| ||| ||||| ||||| ||||| ||||| ||| |||
 Oy 2333 ggcgcagggcacaanaatncttgaactctgggagggcagaggtctgcg tgaagcccaancgc 2382

Db 149032 GGCATTGCATCTCAGGCTTGGGCAACAGAGCAAAAACCTCTGTCCAAAAA 149091
 2383 ggcattgcattccgcgcctgggtaacaagaagaaactgtccaaaaa 2442
 Oy 2383 ggcattgcattccgcgcctgggtaacaagaagaaactgtccaaaaa 2442

RESULT	5			
LOCUS	AF134726	180283 bp	DNA	PRI 27-MAR-1999

```
DEFINITION Homo sapiens BAC clone 215012 NG35, NG36, G9A, NG22, G9, HSP70-2,
HSP70-1, HSP70-HOM, snRNP, G7A, NG37, NG23, and MutsH5 genes,
complete cds.
ACCESSION AF134726
NID 94529886
VERSION AF134726.1 GI:4529886
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 180283)
AUTHORS Rowen,L., Qin,S., Madan,A., Dickhoff,R., Dors,M., Madan,A.,
Hicks,P., Loretz,C., Rattliffe,A., Abbasi,N., Shaffer,T. and
Hood,L.
TITLE Sequence of the human major histocompatibility complex class III
region
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 180283)
AUTHORS Rowen,L.
TITLE Direct Submission
JOURNAL Submitted (15-MAR-1999) Department of Molecular Biotechnology, Box
357730 University of Washington, Seattle, WA 98195, USA
COMMENT Sequencing methodology: high redundancy shotgun in plasmids.
Interspersed Repeats were identified with RepeatMasker (available
from http://ftp.genome.washington.edu/RM/RepeatMasker.html). This
sequence overlaps cosmid S22A (AF019413) by 2363 bases and BAC
210G24 (AF129756) by 12177 bases.
FEATURES
source 1..180283
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="BAC 215012"
/map="6p21.3"
/chromosome="6"
/notes="major histocompatibility complex class III region"
misc_feature 1..2363
/notes="Overlap span with cosmid S22A in AF019413."
repeat_region complement(103..402)
/rpt_type=dispersed
/rpt_family="AluSx"
variation 408..409
/notes="cosmid S22A: ttct; BAC 215012: tt"
/rpt_type=dispersed
/rpt_family="AluSx"
repeat_region complement(420..480)
/rpt_type=dispersed
/rpt_family="L1PA16"
repeat_region complement(458..732)
/rpt_type=dispersed
/rpt_family="L1PA13"
repeat_region complement(834..1134)
/rpt_type=dispersed
/rpt_family="AluSx"
variation 1134..1136
/notes="cosmid S22A: ct; BAC 215012: ctt"
/rpt_type=dispersed
/rpt_family="AluSx"
repeat_region complement(1139..1257)
/rpt_type=dispersed
/rpt_family="AluSg1"
variation 1445..1446
/notes="cosmid S22A: ctt; BAC 215012: ct"
/rpt_type=dispersed
/rpt_family="AluSg1"
repeat_region complement(1446..1746)
/rpt_type=dispersed
/rpt_family="AluSx"
variation 1843
/notes="cosmid S22A: t; BAC 215012: c"
/rpt_type=dispersed
/rpt_family="AluSx"
repeat_region 2430..2489
/rpt_type=dispersed
/rpt_family="MLT1A2"
repeat_region complement(2831..2884)
/rpt_type=dispersed
/rpt_family="LINE2"
repeat_region complement(3091..3392)
/rpt_type=dispersed
/rpt_family="AluY"
repeat_region complement(3416..3712)
/rpt_type=dispersed
/rpt_family="AluSg"
repeat_region complement(3719..3790)
/rpt_type=dispersed
/rpt_family="MLT2FA"
repeat_region 3813..4084
/rpt_type=dispersed
/rpt_family="AluSg"
repeat_region 4086..4375
/rpt_type=dispersed
/rpt_family="AluSg"
repeat_region complement(4417..4729)
/rpt_type=dispersed
/rpt_family="AluSg"
repeat_region 4781..4910
/rpt_type=dispersed
/rpt_family="FLAM_A"
repeat_region complement(4939..4968)
/rpt_type=dispersed
/rpt_family="MLT2E"
repeat_region complement(4980..5112)
/rpt_type=dispersed
/rpt_family="Aluub"
repeat_region complement(5167..5468)
/rpt_type=dispersed
/rpt_family="AluSx"
repeat_region complement(5607..5738)
/rpt_type=dispersed
/rpt_family="LINE2"
repeat_region 5786..5873
/rpt_type=dispersed
/rpt_family="L1ME3A"
repeat_region 5885..5938
/rpt_type=dispersed
/rpt_family="MER5A"
repeat_region 5955..6090
/rpt_type=dispersed
/rpt_family="FLAM_C"
repeat_region 6099..6389
/rpt_type=dispersed
/rpt_family="AluSx"
repeat_region 6403..6446
/rpt_type=dispersed
/rpt_family="MER5A"
repeat_region complement(6711..7011)
/rpt_type=dispersed
/rpt_family="AluSx"
repeat_region complement(7029..7330)
/rpt_type=dispersed
/rpt_family="AluSx"
repeat_region 7486..7688
/rpt_type=dispersed
/rpt_family="Aluub"
repeat_region 7689..7851
/rpt_type=dispersed
/rpt_family="L1MB8"
repeat_region 7852..8150
/rpt_type=dispersed
/rpt_family="AluSg"
repeat_region 8151..8263
/rpt_type=dispersed
/rpt_family="L1MB6"
repeat_region 8290..8425
/rpt_type=dispersed
/rpt_family="FLAM_C"
repeat_region 8451..8730
/rpt_type=dispersed
```



```

repeat_region 3081..3178
/note="Trigger3(golem) repeat: matches 1..98 of consensus"
repeat_region 3177..3243
/note="MUT1A2 repeat: matches 284..351 of consensus"
variation 3673..3675
/note="clone 498124; caa in this entry; substitution"
/note="cgcg"
3774..3776
/note="clone 498124; ata in this entry; substitution"
/note="taca"
4028..4177
/note="L2 repeat: matches 2559..2733 of consensus"
variation 4045..4047
/note="clone 498124; cag in this entry; substitution"
/note="cgg"
4524..4526
/note="clone 498124; ttt in this entry; substitution"
/note="tct"
4651..4653
/note="clone 498124; cct in this entry; substitution"
/note="tct"
4865..4867
/note="clone 498124; gta in this entry; substitution"
/note="gca"
5060..5062
/note="clone 498124; cgt in this entry; substitution"
/note="cat"
5111..5113
/note="clone 498124; att in this entry; substitution"
/note="agt"
5210..5615
/note="L1MD2 repeat: matches 5746..6164 of consensus"
variation 5349..5351
/note="clone 498124; cag in this entry; substitution"
/note="cgg"
5444..5447
/note="clone 498124; actg in this entry; substitution"
/note="atcg"
5588..5590
/note="clone 498124; ttt in this entry; substitution"
/note="tct"
5670..5672
/note="clone 498124; cat in this entry; substitution"
/note="cgt"
5701..5703
/note="clone 498124; gca in this entry; substitution"
/note="gta"
5818..6130
/note="AluSq repeat: matches 1..313 of consensus"
5916..5918
/note="clone 498124; acg in this entry; substitution"
/note="atg"
6211..6213
/note="match: GSSS A0224537 B56744"
6445..6447
/note="clone 498124; gtg in this entry; substitution"
/note="gcg"
6774..6776
/note="clone 498124; cat in this entry; substitution"
/note="cgt"
6836..6838
/note="clone 498124; gtt in this entry; substitution"
/note="gct"
6924..6926
/note="clone 498124; tgc in this entry; substitution"
/note="tac"
6940..6942
/note="clone 498124; aaa in this entry; substitution"
/note="aga"
7288..7291
/note="clone 498124; ttc in this entry; substitution"
/note="tcc"

variation /note="clone 498124; tggt in this entry; substitution"
/note="tcgg"
7478..7480
/note="clone 498124; act in this entry; substitution"
/note="agct"
7823..7825
/note="clone 498124; gtc in this entry; substitution"
/note="gcc"
8324..8326
/note="clone 498124; cgt in this entry; substitution"
/note="cat"
8478..8480
/note="clone 498124; gtc in this entry; substitution"
/note="gcc"
8516..8518
/note="clone 498124; gcc in this entry; substitution"
/note="gtc"
8834..8863
/note="L2 repeat: matches 2714..2743 of consensus"
9029..9030
/note="clone 498124; ca in this entry; deletion"
/note="cagaga"
9030..9077
/note="24 copies 2 mer ag 96% conserved"
9127..9129
/note="clone 498124; cct in this entry; substitution"
/note="ctt"
9160..9162
/note="clone 498124; tcg in this entry; substitution"
/note="ttg"
9303..9608
/note="AluY repeat: matches 1..304 of consensus"
9517..9519
/note="clone 498124; cgg in this entry; substitution"
/note="cag"
9587..9594
/note="clone 498124; aaaaaaa in this entry; insertion"
9610..9821
/note="MIR repeat: matches 34..258 of consensus"
9629..9631
/note="clone 498124; tgc in this entry; substitution"
/note="ttc"
9822..10484
/note="MER6 repeat: matches 1..697 of consensus"
10038..10040
/note="clone 498124; atg in this entry; substitution"
/note="acg"
10056..10058
/note="clone 498124; cat in this entry; substitution"
/note="cgt"
10171..10173
/note="clone 498124; cgt in this entry; substitution"
/note="cat"
10219..10220
/note="clone 498124; gc in this entry; deletion"
/note="gcc"
10485..10764
/note="AluY repeat: matches 1..285 of consensus"
10530..10532
/note="clone 498124; ggg in this entry; substitution"
/note="ggc"
10606..10608
/note="clone 498124; aca in this entry; substitution"
/note="aaa"
10765..10939
/note="MER6 repeat: matches 697..865 of consensus"

Note: remainder of annotations omitted.
Query Match 9.5%; Score 232; DB 29; Length 185371;
Best Local Similarity 87.4%; Pred. No. 5.14e-154;
Matches 264; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

```

Db	62805	TTTTTTTTTTTTTTGAGATGAGATTGGCTGTTCGCCAGCGTGGAGTGCAATGGC	62864
Cp	2442	tttttttttttttttggacagatttcgtcttgtaaccaggctgtagtgcattggc	2383
Db	62865	GTCATCTTGCGCCACCGCACCTCTGCCTCCCGGGGTTCAGCAAAATCTCCTCCATGACC	62924
Cp	2382	gsgnctctgggtccacgcgaactctgcctcccacagtcaaatntcttcgctcgagcc	2323
Db	62925	TCCCTAGTAGTCGGGATTTACTCATGCGCCACACGCCCCGGCTAATTTTGATTTTNG	62984
Cp	2322	tccctagtagtccaagnttacagcactgcgtcacacgcgccgctaattngtatlttng	2263
Db	62985	TAGAGATGGGGTTCCTCCATGTGTGGTCAGAGCTGAGTCTCAACTCCCGACTGAGGTGANC	63044
Cp	2262	tagagatgggggttcttcacgcgttgttcaagctcggtcttgaactccgcgacctagatgac	2203
Db	63045	CACCCACCTCGGCCCCCACAAGTGTCTGAATTACAGATGTAACGACGACGACCTGGCCCA	63104
Cp	2202	caccacaccttgcctccccaagtgttgattacaagttgtgagccactgcgcttgcgcaa	2143
Db	63105	GA 63106	
Cp	2142	ga 2141	
RESULT	7		
LOCUS	HSB10B1	40926 bp	DNA PRI 17-AUG-1998
DEFINITION	Human DNA sequence from cosmid B10B1 on chromosome 22 Contains ESTs, CA repeat and STS, complete sequence.		
ACCESSION	Z73979		
NID	91403046		
VERSION	Z73979.1	GI:1403046	
KEYWORDS	HFG.		
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;		
	Primates; Catarrhini; Hominoidea; Homo.		
REFERENCE	1 (bases 1 to 40926)		
AUTHORS	Hunt,S.		
JOURNAL	Direct Submission		
COMMENT	Submitted (05-JUN-1996) E-mail enquiries: humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk IMPORTANT: This sequence is the entire insert of clone B10B1. The true left end of clone B10B1 is at 1 in this sequence. The true right end of clone B10B1 is at 40926. B10B1 is from the human chromosome 22-specific cosmid library (SC22CB) constructed at the Sanger Centre by Mark Ross and Cordellia Langford. This sequence was generated from part of bacterial clone contigs of human chromosome 22, constructed by the Sanger Centre chromosome 22 mapping group. Further information can be found at http://www.sanger.ac.uk/HGP/Chr22/ .		
FEATURES	Location/Qualifiers		
source	1..40926		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="22"		
	/map="q12"		
	/clone="B10B1"		
	/clone_1bp="SC22CB"		
repeat_region	2247..2369		
	/note="MIR element fragment"		
misc_feature	2517..2795		
	/note="match: EST T07591 clone HPEBL36"		
repeat_region	5618..5672		
	/note="MIR element fragment"		
misc_feature	8057..8385		
	/note="match: Z17028 DNA segment containing (CA) repeat"		
repeat_region	8158..8207		
	/note="25 copies of AC 100 & conserved; Polymorphic by comparison with Z17028"		
repeat_region	9513..9575		

			/note="MER3 element fragment"	
	repeat_region	9674 . 9749	/note="MER3 element fragment"	
	repeat_region	11611 . 11752	/partial	
	repeat_region	/note="Alu repeat: matches 1. .145 of consensus"	/partial	
	repeat_region	11762 . 11898	/note="Alu repeat: matches 171. .308 of consensus"	
	repeat_region	16545 . 16836	/note="Alu repeat: matches 1. .308 of consensus"	
	repeat_region	17577 . 17609	/note="11 copies of 3 mer 88 % conserved"	
	repeat_region	18668 . 18958	/note="Alu repeat: matches 1. .308 of consensus"	
	repeat_region	20051 . 20338	/note="Alu repeat: matches 1. .308 of consensus"	
	repeat_region	20339 . 20368	/note="15 copies of 2 mer 100 % conserved"	
	misc_feature	20374 . 20545	/note="match: 5' EST D31518"	
	repeat_region	20819 . 21108	/note="Alu repeat: matches 1. .308 of consensus"	
	misc_feature	23882 . 24398	/note="match: multiple ESTs"	
	repeat_region	27345 . 27374	/note="6 copies of 5 mer 97 % conserved"	
	repeat_region	27409 . 27702	/partial	
	repeat_region	28831 . 29133	/note="Alu repeat: matches 308. .1 of consensus"	
	repeat_region	/partial	/note="Alu repeat: matches 308. .1 of consensus"	
	repeat_region	30874 . 30968	/note="MIR2 element fragment"	
	repeat_region	30917 . 30953	/note="MIR element fragment"	
	misc_feature	complement(32935..33185)	/note="match: STS G03723"	
	repeat_region	35799 . 35903	/note="MIR2 element fragment"	
	repeat_region	36321 . 36525	/note="L1 element fragment"	
	unsure	36368 . 36371	/note="Could be 5 G's"	
	repeat_region	39290 . 39700	/partial	
	repeat_region	/note="Alu repeat: matches 308. .1 of consensus"	/partial	
	BASE COUNT	10681 a 9179 c 9923 g 11143 t		
	ORIGIN			
D	Query Match	9.4% Score 229; DB 29; Length 40926;		
Bd	Best Local Similarity 86.6%; Pred.No.1,36e-151;			
Oy	Matches 264; Conservative 0; Mismatches 41; Indels 0; Gaps 0;			
Db	16538 AAGTTGAGCGCCGACCGCAGTCCTGTATCTCAAGCACTTGGGAGGCCAGGC	16597		
Oy		2138 aagcttcgccaaggcagggtgcacacctgtatcccaaccattgggaagccaaagt		
Db	16598 AGCGGATCACACTAGGTCGGGATTGGAACCAAGCTTGACAATGAGAACCOCAT	16657		
Oy		2198 gggtgatcatcctaagcccggttcaaacacagccttgaccaactggaagaacccat		
Db	16658 CTCTACTAAATAATTAAGCCGGCGCTGGTGGCGCATGCCGTGATCCAGCTACT	16717		
Oy		2258 ctctactnaaatcnaatttagccggtgltagtcgcatgctgtaancctgtctact		
Db	16718 CGGGAGGCTGAGCGAGGAATGCTTGACACCTGGAGCGAGAGTTGCAATAGCCGAG	16777		
Oy		2318 caggagggcaggaggaanaattctttgaacttgggagggagcaggttgcgtgagccag		

D	b	16778	ATCGACCATTCGACTCCGAGCTGGGCAACAGACAATAAATCATTCTCAAAAAAA	16837
Oy		2378	atcgccattgctcactccagctgggttaacaaagcaaacctctgccaaaaaaa	2437
Db	16838	AAAAA	16842	
Oy		2438	aaaaa	2442
RESULT	8			
LOCUS		AC006968	96115 bp	DNA HTG 05-MAR-1999
DEFINITION		Homo sapiens clone Df0649M07,	WORKING DRAFT SEQUENCE,	2 unordered pieces.
ACCESSION		AC006968		
NTD		94337292		
VERSION		AC006968.1	GI:4337292	
KEYWORDS		HTG; HTGS_PHASEL.		
SOURCE		human.		
- ORGANISM		Homo sapiens		
REFERENCE		Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;		
AUTHORS		Primates; Catarrhini; Hominiidae; Homo.		
JOURNAL		1 (bases 1 to 96115)		
REFERENCE		Waterston,R.H.		
AUTHORS		The sequence of Homo sapiens clone		
TITLE		unpublished		
JOURNAL		2 (bases 1 to 96115)		
REFERENCE		Waterston,R.H.		
AUTHORS		Direct Submission		
TITLE		Submitted (05-MAR-1999) Genome Sequencing Center, Washington		
JOURNAL		University School of Medicine, 4444 Forest Park Parkway, St. Louis,		
		MO 63108, USA		
COMMENT		* NOTE: This is a 'working draft' sequence. It currently		
		* consists of 2 contigs. The true order of the pieces		
		* is not known and their order in this sequence record is		
		* arbitrary. Gaps between the contigs are represented as		
		* runs of N, but the exact sizes of the gaps are unknown.		
		* This record will be updated with the finished sequence		
		* as soon as it is available and the accession number will		
		* be preserved.		
	*	1 21676; contig of 21676 bp in length		
	*	21677 21694; gap of unknown length		
	*	21695 96113; contig of 74421 bp in length.		
FEATURES		Location/Qualifiers		
source		1..96115		
		/organism="Homo sapiens"		
		/db_xref="taxon:9606"		
		/clone="Df0649M07"		
BASE COUNT		32834 a 18141 c 17098 g 28024 t	18 others	
ORIGIN				
Query Match		9..3%; Score 226; DB 20; Length 96115;		
Best Local Similarity		86..9%; Pred. No. 3 60e-149;		
Matches	259; Conservative	0; Mismatches 39; Indels 0; Gaps 0;		
Db	68620	TTTTTTTTTTTTTTTACTGTGAGTTTCGCTTGTTGCCAGGCTGAGTAGATGGC	68679	
Cp	2442			
		ttttttttttttttttagcacagtgttcgtctcttgtaaccaggctgagagcatgc	2383	
Db	68680	ATGATCTCGGCTCATCTGCACACTCTGCTCCCTGGGTTTAAGAATTTCTCTCCACG	68739	
Cp	2382			
		gcgacctgagccacacgaacctgcgcctcccacaattcaagaattntctcctcgcc	2323	
Db	68740	TCCCTAGTAGCTAGGATTTCACAGCATGTGSCCACACAGCCCGGCTAATTTTGATTTGAG	68799	
Cp	2322	tctctagtagccaagnttacacgatgcctacgcctacacagcccgctaattngtatltmag	2263	
Db	68800	TAGAGACGGGGTTTCTCATGTGTGCTCAGCGTGGTCTCGAACTCCCGACCTCAGATGATC	68859	
Cp	2262	tagagataggggttctccaccgltgtgcacagctgcttgttaactcccgcgcttagatgc	2203	
Db	68860	CACCCACCTCGGCGCCCCAAAGTGGCGGGATTACAGGCGTGAAGCACACACCGCCGGC	68917	

```

cp 2202 caccactctgctcccaagtgcttggtatcacagtgtagcaccatcgccctggcc 2145

RESULT 9
LOCUS HS302D9 145871 bp DNA PRI 24-NOV-1998
DEFINITION Human DNA sequence from PAC 302D9 on chromosome 22q11.2-qter.
          Contains STS, complete sequence.
ACCESSION 282198
NID 93164067
VERSION 262198.1
KEYWORDS GI:3164067
SOURCE HTG.
ORGANISM human.
          Homo sapiens
          Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
          Primates; Catarrhini; Hominiinae; Homidae; Homo.
REFERENCE 1 (bases 1 to 145871)
AUTHORS Bridgeman,A.
TITLE Direct Submission
JOURNAL Submitted (02-JUL-1997) E-mail enquires: humquery@sanger.ac.uk

COMMENT
Clone requests: clonerequests@sanger.ac.uk
On May 28, 1998 this sequence version replaced gi:1772948.
IMPORTANT: This sequence is the entire insert of clone 302D9.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variations annotated may not be found in the sequence submission
corresponding to the overlapping clone as we submit sequences with
only a small overlap. This sequence was generated from part of
bacterial clone contigs of human chromosome 22, constructed by the
Sanger Centre chromosome 22 mapping group. Further information can
be found at http://www.sanger.ac.uk/HGP/Chr22/
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu). Where the
sequence is ambiguous, there is an annotation using the 'unsure'
feature key.
The true left end of clone 302D9 is at 1 in this sequence. The true
right end of clone 41562 is at 35164.
The true left end of clone 282F2 is at 69679.
The true right end of clone 302D9 is at 145871.
302D9 is from the library RPcII constructed at the Roswell Park
Cancer Institute by the group of Pieter de Jong.
For further details see http://bacpac.med.buffalo.edu/.

FEATURES
     source
         1..145871
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="22"
            /map="q11.2-qter"
            /clone="302D9"
            /clone_id="RPcII"
            188..244
            /note="MER3 repeat: matches 209. .155 of consensus"
            246..561
            /note="Alusx repeat: matches 1. .302 of consensus"
            585..759
            /note="MER3 repeat: matches 166. .1 of consensus"
            889..933
            /note="MER5A repeat: matches 71. .26 of consensus"
            1029..1336
            /note="Alus repeat: matches 303. .1 of consensus"
            1446..1583
            /note="MIR repeat: matches 20. .160 of consensus"
            2359..2659
            /note="Alusq repeat: matches 299. .3 of consensus"
            2679..2980
            /note="Alus repeat: matches 302. .2 of consensus"
            3344..3643
            /note="Alu repeat: matches 2. .301 of consensus"
            3747..3922
            /note="MIR repeat: matches 188. .365 of consensus"
            3926..4275

```

repeat_region	/note="M1E1B repeat: matches 3. .364 of consensus" 4296. .4482
repeat_region	/note="M1E1E repeat: matches 382. .568 of consensus" 5188. .5488
repeat_region	/note="A1uSg repeat: matches 299. .1 of consensus" 6366. .6447
repeat_region	/note="M1R2 repeat: matches 141. .59 of consensus" 6644. .6680
repeat_region	/note="MADE1 repeat: matches 1. .37 of consensus" 6683. .6584
repeat_region	/note="A1uSx repeat: matches 302. .1 of consensus" 6988. .7036
repeat_region	/note="MADE1 repeat: matches 33. .80 of consensus" 7479. .7751
repeat_region	/note="A1uDo repeat: matches 9. .290 of consensus" 7773. .8057
repeat_region	/note="A1uDo repeat: matches 2. .295 of consensus" 8407. .8473
repeat_region	/note="M1R2 repeat: matches 146. .79 of consensus" 8911. .9027
repeat_region	/note="M1R repeat: matches 262. .147 of consensus" 9107. .9277
repeat_region	/note="M1R repeat: matches 91. .262 of consensus" 9270. .9409
repeat_region	/note="M1R repeat: matches 155. .15 of consensus" 9518. .9676
repeat_region	/note="FM repeat: matches 161. .3 of consensus" 9817. .10222
variation	/note="M1TA repeat: matches 2. .425 of consensus" 10023. .10025
variation	/note="CTG in clone 415G2" /replace="ctg" 10309. .10377
repeat_region	/note="M1R repeat: matches 150. .82 of consensus" 11835. .11943
repeat_region	/note="M1R repeat: matches 20. .137 of consensus" 12488. .12660
repeat_region	/note="M1R repeat: matches 97. .262 of consensus" 13198. .13200
variation	/note="CGT in clone 415G2" /replace="cgt" 13395. .13685
repeat_region	/note="A1uSg repeat: matches 302. .1 of consensus" 13697. .13807
repeat_region	/note="M1R repeat: matches 191. .76 of consensus" 13824. .13916
repeat_region	/note="M1R repeat: matches 91. .189 of consensus" 13947. .14057
repeat_region	/note="M1R repeat: matches 29. .146 of consensus" 14058. .14364
repeat_region	/note="A1uSg repeat: matches 300. .1 of consensus" 14367. .14449
repeat_region	/note="M1R repeat: matches 137. .225 of consensus" 14865. .15037
repeat_region	/note="M1R repeat: matches 49. .233 of consensus" 15301. .15396
repeat_region	/note="M1B1B repeat: matches 1. .99 of consensus" 15486. .15601
repeat_region	/note="A1uSg repeat: matches 1. .116 of consensus; incomplete repeat" 15666. .15774
repeat_region	/note="M1B1B repeat: matches 119. .175 of consensus" 15725. .16024
repeat_region	/note="A1uSg repeat: matches 1. .299 of consensus" 16041. .16243
variation	/note="M1B1B repeat: matches 188. .390 of consensus" 16309. .16311
repeat_region	/note="CTG in clone 415G2" /replace="ctg" 16544. .16851
repeat_region	/note="A1uY repeat: matches 300. .2 of consensus" 18391. .18701
repeat_region	/note="A1uB repeat: matches 1. .302 of consensus"

[illegible]

ORGANISM Homo sapiens
Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Mammalia;
Eutheria: Primates: Catarrhini: Homiidae: Homo.
1 (bases 1 to 160258)
Mashreghi-Mohammadi, M.
AUTHORS
TITLE Direct Submission
JOURNAL Submitted (09-MAR-1999) Wellcome Trust Genome Campus, Hinxton,

COMMENT

represent the correct sequence. Work on the sequence is in progress and the release of this data is based on the understanding that the sequence may change as work continues. The sequence may be contaminated with foreign sequence from *E. coli*, yeast, vector, phage etc. Order of segments is not known: 800 n's separate segments. Unfinished: dj5000L14 Contig_ID: 02706 acc=AL023583 Length: 20402 bp Unfinished: dj5000L14 Contig_ID: 00165 acc=AL023583 Length: 10585 bp Unfinished: dj5000L14 Contig_ID: 01976 acc=AL023583 Length: 28901 bp Unfinished: dj5000L14 Contig_ID: 01917 acc=AL023583 Length: 1050 bp Unfinished: dj5000L14 Contig_ID: 01722 acc=AL023583 Length: 1585 bp Unfinished: dj5000L14 Contig_ID: 02443 acc=AL023583 Length: 1424 bp Unfinished: dj5000L14 Contig_ID: 02540 acc=AL023583 Length: 1014 bp Unfinished: dj5000L14 Contig_ID: 01799 acc=AL023583 Length: 1244 bp Unfinished: dj5000L14 Contig_ID: 01123 acc=AL023583 Length: 4745 bp Unfinished: dj5000L14 Contig_ID: 03094 acc=AL023583 Length: 5968 bp Unfinished: dj5000L14 Contig_ID: 02910 acc=AL023583 Length: 18171 bp Unfinished: dj5000L14 Contig_ID: 03010 acc=AL023583 Length: 2469 bp.

* NOTE: This is a "working draft" sequence.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

FEATURES

Location/Qualifiers

Source

[illegible]

Query Match	9.3%;	Score 226;	DB 19;	Length 160258;
Best Local Similarity	86.2%;	Pred. No. 3.60e-149;		

Matches	262;	Conservative	0;	Mismatches	42;	Indels	0;	Gaps	0
---------	------	--------------	----	------------	-----	--------	----	------	---

[illegible]

Db	29718	GGGAGATCTCTTGAGCGCTGGGAGTTCAAGACCAAGCTGACCAACATGSGAAGAACCCCGT	29777
Qy	2198	gggtgtgatacactcaagccgggagltcaagaacacgctgcagcaacgctgsgaatacccat	2257
Db	29778	CTCTACTAAAAATACAAAATTATGCAAGCGGTGTGGCGCATGCTGTATTCACGACTACT	29833
Qy	2258	ctctactcaaaaatacacaattagccgggctggtgtagcgcactgctgtgaatacccgtgctact	2317
Db	29838	TGGGAGGCTTAAGGACAGAGAAATTGGCTTGAAACCCAGAGAGCGGAGCTGTGTGTAGCCGAG	29897
Qy	2318	caggagagccgaggcagaaaataattccttgaactgtgggagccagagtgctgcggtgaagcccg	2377
Db	29898	ATCATGCGCATTTGCACATCCGGCTGGGCAAGAGACAAACTCCATCTCAAAAAAAAAAAAAA	29957
Qy	2378	ancgcgcattgcactccagcctgtggttaacaagagcaaaactctgtctcaaaaaaaaaaaaaa	2437
Db	29958	AAAA 29961	
Qy	2438	aaaa 2441	

LOCUS	DEFINITION	PR1	28-JAN-1999
HS821D11	Human DNA sequence from clone 821D11 on chromosome 22q12.3-13.1. Contains three partial unknown genes, one downstream of a predicted CpG island, and the first coding exon of the SREBF2 gene for Sterol Regulatory Element Binding Transcription Factor 2 downstream of a predicted CpG island. Contains ESTs, STS, GSSs and genomic marker D2S31157, complete sequence.		

KEYWORDS HTG; CpG island; D22S1157; SREBF2; SREBP2; Sterol Regulatory Element Binding Transcription Factor 2.

SOURCE	human.
ORGANISM	Homo sapiens

REFERENCE
1. (bases 1 to 76727)
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria
Primates; Catarrhini; Hominae; Homo.
1 (bases 1 to 76727)

JOURNAL Submitted (11-MVG-1998) E-mail enquiries: hunquerry@sanger.ac.uk
 Clone requests: clonerequest@sanger.ac.uk
 On Aug 12, 1998 this sequence version replaced gi:3355580.

COMMENT

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

IMPORTANT: This sequence is not the entire insert of clone 82D11. It may be shorter because we only sequence overlapping sections once, or longer because we arrange for a small overlap between neighbouring submissions.

The true left end of clone 250B10 (2291196) is at 76828 in this sequence. The true right end of clone 10936 (AL023879) is at 42082. This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the annotation is ambiguous, there is an annotation using the 'unsure' feature key.

This sequence was generated from part of bacterial clone contigs of human chromosome 22, constructed by the Sanger Centre Chromosome Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr22>

82BD1 is from the library RPC15 constructed at the Roswell Park Cancer Institute by the group of Peter de Jong. For further details see <http://bacpac.med.buffalo.edu/VECTOR:PCPAC2>.

```
FEATURES      Location/Qualifiers
source        1. .76727
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
```

```

repeat_region /chromosome="22"
               /clone="821D11"
               /map="q12.3-13.1"
               /clone.lib="RPC15"
               complement(445..743)
               /note="Aluub repeat: matches 302. .8 of consensus"
               446..448
               /note="clone 10966; tct in this entry; substitution"
               /replace="ttt"
               /gene="dJ821D11.1"
               join(1230..1711,13695..>13736)
               /note="this gene and dJ821D11.2 could be part of one gene;
               match: ESTs AA19437 AA332515"
               /evidence="not_experimental"
               /product="dJ821D11.1 (PUTATIVE protein)"
               1230..13736
               /gene="dJ821D11.1"
               1314..1316
               /note="dJ821D11.1"
               /note="clone 10966; tgt in this entry; substitution"
               /replace="tct"
               join(1601..1711,13695..>13736)
               /gene="dJ821D11.1"
               /note="other possible startcodon at 1583"
               /codon_start=1
               /evidence="not_experimental"
               /product="dJ821D11.1 (PUTATIVE protein)"
               /protein_id="CA16279.1"
               /db_xref="PID:94200330"
               /db_xref="GI:4200330"
               /translation="MALTLAKADSPRTALCSNMLTASFSAQGHKSLQKPLLSQ
               ACVGCLE"
               1760..1762
               /gene="dJ821D11.1"
               /note="clone 10966; gag in this entry; substitution"
               /replace="ggg"
               1773..1895
               /note="MIR repeat: matches 47. .182 of consensus"
               1901..1903
               /gene="dJ821D11.1"
               /note="clone 10966; tca in this entry; substitution"
               /replace="tta"
               2302..2304
               /gene="dJ821D11.1"
               /note="clone 10966; gga in this entry; substitution"
               /replace="gaa"
               complement(3042..3133)
               /note="LIMC1 repeat: matches 1079. .988 of consensus"
               complement(3162..3456)
               /note="AluSp repeat: matches 297. .2 of consensus"
               complement(3457..3551)
               /note="MIR repeat: matches 149. .59 of consensus"
               complement(3552..3847)
               /note="Aluub repeat: matches 299. .2 of consensus"
               4753..4754
               /gene="dJ821D11.1"
               /note="clone 10966; tt in this entry; deletion"
               /replace="tttgt"
               4791..4792
               /gene="dJ821D11.1"
               /note="clone 10966; tt in this entry; deletion"
               /replace="ttattatt"
               complement(4800..5099)
               /note="AluSg repeat: matches 303. .2 of consensus"
               5107..5112
               /gene="dJ821D11.1"
               /note="clone 10966; ttatt in this entry; insertion"
               /replace="ttt"
               complement(5128..5427)
               /note="AluY repeat: matches 300. .1 of consensus"
               5359..5361
               /gene="dJ821D11.1"

variation
repeat_region /note="clone 10966; att in this entry; substitution"
               /replace="act"
               complement(5430..5729)
               /note="AluSg repeat: matches 300. .1 of consensus"
               5730..5775
               /note="23 copies 2 mer tt 87% conserved"
               5761..5762
               /gene="dJ821D11.1"
               /note="clone 10966; at in this entry; deletion"
               /replace="att"
               5772..5773
               /gene="dJ821D11.1"
               /note="clone 10966; tt in this entry; deletion"
               /replace="tat"
               complement(5777..6075)
               /note="AluY repeat: matches 299. .1 of consensus"
               5778..5779
               /gene="dJ821D11.1"
               /note="clone 10966; tt in this entry; deletion"
               /replace="ttat"
               5780..5782
               /gene="dJ821D11.1"
               /note="clone 10966; tat in this entry; substitution"
               /replace="ttt"
               5799..5801
               /gene="dJ821D11.1"
               /note="clone 10966; cgg in this entry; substitution"
               /replace="cgg"
               5868..5870
               /gene="dJ821D11.1"
               /note="clone 10966; ccg in this entry; substitution"
               /replace="ctg"
               complement(6080..6211)
               /note="FLM.C repeat: matches 133. .1 of consensus"
               6432..6434
               /gene="dJ821D11.1"
               /note="clone 10966; atg in this entry; substitution"
               /replace="agg"
               complement(7157..7269)
               /note="MER42c repeat: matches 1538. .1419 of consensus"
               7588..7653
               /note="MIR2 repeat: matches 79. .145 of consensus"
               7748..7889
               /note="MER46 repeat: matches 88. .232 of consensus"
               complement(7895..8177)
               /note="AluSx repeat: matches 302. .18 of consensus"
               complement(8307..8606)
               /note="AluSx repeat: matches 300. .1 of consensus"
               complement(8708..8887)
               /note="Aluub repeat: matches 297. .117 of consensus"
               8898..9197
               /note="AluSx repeat: matches 1. .302 of consensus"
               9047..9049
               /gene="dJ821D11.1"
               /note="clone 10966; gcg in this entry; substitution"
               /replace="gtg"
               complement(9207..9329)
               /note="Aluud repeat: matches 123. .1 of consensus"
               complement(9330..9609)
               /note="LIMB3 repeat: matches 930. .640 of consensus"
               complement(9948..10250)
               /note="AluSx repeat: matches 301. .1 of consensus"
               complement(10457..10754)
               /note="AluSg repeat: matches 300. .1 of consensus"
               10699..10701
               /gene="dJ821D11.1"
               /note="clone 10966; cgc in this entry; substitution"
               /replace="cac"
               10876..10878
               /gene="dJ821D11.1"
               /note="clone 10966; ccg in this entry; substitution"
               /replace="ctg"
               10969..10971

```

/gene="dJ821D11.1"
/note="clone 109G6; ggc in this entry; insertion"
/replace="gc"

*** remainder of annotations omitted.

Query Match 9.2%; Score 223; DB 29; Length 76727;

Best Local Similarity 87.8%; Pred. No. 9,48e-147;

Matches 232; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Db 57454 GTGGCTCACACCTTAATCCAGACCTTTGGAGCGCGGAGGATACCTGAGT 57513

2156 gtggctcacacctgtaatcccaacttggagcgcaagtgggagatcactaaggc 2215

Db 57514 TGGAGTTAGACACCTGACACACATGAGAAACCCGCTCTACTTAAATAACAA 57573

2216 cggagagttacagaccagctgaccacagctgagaaacccactctactnaaatcnaa 2275

Db 57574 ATTGCGCGGCGTCTGGACATGCGCTGTAATCCAGCTCTAGAGAGGTAAAGCAGGA 57633

2276 attagccgagcgtgtagcgcatgctgctactaagagagcgagagagaa 2335

Db 57634 GAATTGCTTAACCCGAGGAGAGAGTTGACAGTGCAGATCGCGCATTCGACTCC 57693

2336 naatttctgaactgggagagcgaggttcggtgagccagangcgccattgcactcc 2395

Db 57694 AGCCTGGGCAACAGACAGACTGTCTCAAAAAA 57740

2396 agcctgggtaacaagcaaacctctgccaaaaa 2442

RESULT 12 HS469D22 85285 bp DNA HTG 16-MAR-1999

LOCUS Human DNA sequence *** SEQUENCING IN PROGRESS *** from clone

469D22, WORKING DRAFT SEQUENCE.

ACCESSION AL031284

VERSION 94455561

KEYWORDS AL031284.2 GI:4455561

SOURCE HTG; HTGS_PHASE1.

ORGANISM human.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 85285)

AUTHORS Cobley, V.

TITLE Direct Submission

JOURNAL Submitted (16-MAR-1999) Wellcome Trust Genome Camps, Hinxton,

Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk

COMMENT On Mar 21, 1999 this sequence version replaced g1:4375971.

IMPORTANT: This sequence is unfinished and does not necessarily

represent the correct sequence. Work on the sequence is in progress

and the release of this data is based on the understanding that the

sequence may change as work continues. The sequence may be

contaminated with foreign sequence from E.coli, yeast, vector,

phage etc. Order of segments is not known: 800 n's separate

segments. Unfinished: dJ469D22 Contig_ID: 01101 acc=AL031284

Length: 61683 bp Unfinished: dJ469D22 Contig_ID: 01792

acc=AL031284 Length: 22802 bp

* NOTE: This is a 'working draft' sequence.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

FEATURES Location/Qualifiers

source

1..85285

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="1"

/clone="469D22"

BASE COUNT 22615 a 19563 c 18556 g 23751 t 800 others

ORIGIN

Query Match 9.2%; Score 224; DB 19; Length 85285;

Best Local Similarity 86.3%; Pred. No. 1.48e-147;

Matches 259; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

Db 2921 TTGGCGGCGACAGTGGCTTAATGCTTAATCCAGACCTTTGGAGGCGAAGTGGCG 2980

2143 ttggcgagcgagcagtggtcactacactgtatcccaacttggagcgcaagtggtg 2202

Db 2981 GATCAGCTGAGGTGAGGAGTTGAGACGACGATGACCAACATGAGAAACCTGTCTTA 3040

2203 gatcatcgaagcgaggttcaagaccagctgacacagctgagaaacccactctta 2262

Db 3041 CTAATAATACAAATTACCGCGCGTGTGTCATGCTTAATCCAGCTACTACAGA 3100

2263 cttaaatatcaaatatgacgagcggtgtagcagatgctgaanccctgactactcaga 2322

Db 3101 GCGTGGAGGAGAGATGATGACACCGAGGAGGAGTGGTGGTGGAGATAGC 3160

2322 ggcgagcgagcaanaatncttgaacttggagcgagaggttcggtgagccagancgc 2382

Db 3161 GCCATTGCACTTCAGCGTGGGCAACAGAGCAAAACTCGTCTTAAAAAAGAA 3220

2383 gccattgcacttcagcgtggtaacaagagcaaacctctgccaaaaa 2442

RESULT 13 AC005087 105984 bp DNA HTG 12-JUN-1998

LOCUS Homo sapiens clone RG306D19, WORKING DRAFT SEQUENCE, 17 unordered

pieces.

ACCESSION AC005087

VERSION 93212904

KEYWORDS AC005087.1 GI:3212904

SOURCE HTG; HTGS_PHASE1.

ORGANISM human.

Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;

Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 105984)

AUTHORS Waterston, R.H.

TITLE The sequence of Homo sapiens clone

Unpublished

JOURNAL 2 (bases 1 to 105984)

Waterston, R.H.

COMMENT Direct Submission

Submitted (12-JUN-1998) Genome Sequencing Center, Washington

University School of Medicine, 4444 Forest Park Parkway, St. Louis,

MO 63108, USA

* NOTE: This is a 'working draft' sequence. It currently

* consists of 17 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

1 1848: contig of 1848 bp in length

* 1849: gap of unknown length

* 1866: contig of 3138 bp in length

* 5004: gap of unknown length

* 5021: contig of 2285 bp in length

* 7306: gap of unknown length

* 7322: gap of unknown length

* 10300: contig of 2977 bp in length

* 10317: gap of unknown length

* 10317: contig of 2573 bp in length

* 12889: gap of unknown length

* 12907: gap of unknown length

* 12907: contig of 2744 bp in length

* 15651: gap of unknown length

* 15668: contig of 3842 bp in length

* 19510: gap of unknown length

* 19527: contig of 3429 bp in length

* 22973: gap of unknown length

* 22973: contig of 3402 bp in length

* 26375: gap of unknown length

* 26392: contig of 4547 bp in length

repeat_region	/note="low coverage" /rpt_family="Aluo" complement(2441. ,2725)
repeat_region	/rpt_family="Aluo" complement(2576. ,3174)
repeat_region	/rpt_family="MER20" complement(3315. ,3385)
repeat_region	/rpt_family="Alusq" complement(3389. ,3695)
repeat_region	/rpt_family="Alusq" complement(3696. ,3831)
repeat_region	/rpt_family="Alusq" complement(4420. ,4522)
repeat_region	/rpt_family="Alusx" 4600. ,4897
repeat_region	/rpt_family="Aluo" 4911. ,5208
repeat_region	/rpt_family="Alu" 5241. ,5622
repeat_region	/rpt_family="L1" complement(5704. ,6003)
repeat_region	/rpt_family="Alusx" 6551. ,6829
repeat_region	/rpt_family="Alusg" complement(7122. ,7330)
repeat_region	/rpt_family="Aluo" 7636. ,7937
repeat_region	/rpt_family="Alusx" complement(8542. ,8842)
repeat_region	/rpt_family="Alusg" complement(8874. ,9007)
repeat_region	/rpt_family="FLAN_C" 9642. ,9769
repeat_region	/rpt_family="MIR" complement(9779. ,10261)
repeat_region	/rpt_family="MTRIG" complement(10319. ,10604)
repeat_region	/rpt_family="Aluo" complement(10617. ,11083)
repeat_region	/rpt_family="L2" 11354. ,11552
repeat_region	/rpt_family="Alusg" 11956. ,12259
repeat_region	/rpt_family="Alub" 12290. ,12580
repeat_region	/rpt_family="Alub" complement(12598. ,12878)
repeat_region	/rpt_family="Alu" complement(12881. ,13182)
repeat_region	/rpt_family="Alusg" 13183. ,13270
repeat_region	/rpt_family="MIR" complement(14166. ,14334)
repeat_region	/rpt_family="MIR" 14462. ,14746
repeat_region	/rpt_family="Alusx" complement(14834. ,14966)
repeat_region	/rpt_family="Alub" 14967. ,15038
repeat_region	/rpt_family="MIR" complement(15782. ,17681)
repeat_region	/rpt_family="L2" 17852. ,17996
repeat_region	/rpt_family="MIR" 18161. ,18221
repeat_region	/rpt_family="Alu" complement(18249. ,18468)
repeat_region	/rpt_family="Alusg" complement(18483. ,18576)
repeat_region	/rpt_family="Alusg" 18606. ,18770
repeat_region	/rpt_family="RRAM" 18785. ,19082
repeat_region	/rpt_family="Alusg"

repeat_region	complement(19331, .19628)
repeat_region	/rpt_family="AluX"
repeat_region	complement(20059, .20331)
misc_feature	/rpt_family="AluUo"
misc_feature	20089, .20299
repeat_region	/note="low coverage"
repeat_region	20376, .20508
repeat_region	/rpt_family="AluUb"
repeat_region	20516, .20802
repeat_region	/rpt_family="AluX"
repeat_region	20803, .20981
repeat_region	/rpt_family="AluUb"
repeat_region	complement(21287, .21581)
repeat_region	/rpt_family="AluSg"
repeat_region	complement(21588, .21622)
repeat_region	/rpt_family="MIR"
repeat_region	complement(22010, .22164)
repeat_region	/rpt_family="AluUo/FRM"
repeat_region	complement(22458, .22743)
repeat_region	/rpt_family="AluUo"
repeat_region	complement(22771, .22889)
repeat_region	/rpt_family="FLM_C"
repeat_region	23272, .23442
repeat_region	/rpt_family="MIR"
repeat_region	24417, .24457
repeat_region	/rpt_family="GC_rich"
repeat_region	complement(24564, .24588)
repeat_region	/rpt_family="GC_rich"
repeat_region	24653, .24715
repeat_region	/rpt_family="GC_rich"
repeat_region	complement(25157, .25456)
repeat_region	/rpt_family="AluY"
repeat_region	complement(25773, .26057)

Note: remainder of annotations omitted

Query Match	Score	DB	Length
9.2%	224	31	114842

Best Local Similarity	84.9%;	Pred.No. 1.48e-14/;
Matches	265; Conservative	0; Mismatches 47; Indels 0; Gaps 0;

99162 AAAAATTGTAATCTGGCCAGGTGTGGTGCCTACTCTGTAATCCACACTATGGGA 99221

2129 ataactccaagtcttggccagcgcgagtgctcaacacctgtaatcccaactttgga 2188

b 99222 GCCAAGGCGGTCACCTGAGTTGCGAGTTCGAGACCAAGCCTGACCAACATGGAG 99281

2189 ggccaagtggtgcatcatctaagcccgaggttcaagacacagcctgaccaacgttgag 2248

99282 AAACCCACCTTACTAAAAATACAAATTTAGCCGGGCATGCTGGCCATGCCTGTAATC 99341

2249 aaaccatcttactnaaataacnaattagccgycgtgtgtagcgcatgctgttaac 2308

93342 CCAGCTACTCAGGAGGCTGAGGCAGAGAAATCGCTTGAACCCGGGAGGCAGAGGTTGCAG 99401

2309 ctgctactcagagagccgaggcaganaatttcttgaactggggagagcagaggttgccg 2368

99402 TGAGCCAGATGCTGCCATTGCACCTCCAGCCCTGGCAACACAGAGCAAAAACCTCCATCTCAA 99461

2369 tgaagccagancgcgccattgcactccagcctggtatacaagagcaaaactctgtccaa 2428

b 99462 AAAACAACAAA 99473

```

2429          ||| ||| |||
2430          ||| ||| |||
2431          ||| ||| |||
2432          ||| ||| |||
2433          ||| ||| |||
2434          ||| ||| |||
2435          ||| ||| |||
2436          ||| ||| |||
2437          ||| ||| |||
2438          ||| ||| |||
2439          ||| ||| |||
2440          ||| ||| |||

```

ESCU1	LOCUS	125515 bp	DNA	PRI	05-AUG-1998
ESCU1	HS230G1	125515 bp	DNA	PRI	05-AUG-1998

Contains EST, STS and GSS, complete sequence.

ACCESSION 204400
UID 93319673
78445661
CT: 3310573

EXTENSION 264400.1 01.03.2007
KEYWORDS HTG; Xp11.3.

```

SOURCE      human.
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 125515)
AUTHORS     Wray, P.
TITLE       Direct Submissions
JOURNAL     Submitted (08-JUL-1998) Chromosome X Project Group
            (http://www.sanger.ac.uk/HGP/ChrX/) Sanger Centre, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humguy@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
            On Jul 15, 1998 this sequence version: clonerequests@sanger.ac.uk
            IMPORTANT: This sequence is the entire insert of clone 230G1.
            During sequence assembly data are compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variations annotated may not be found in the sequence submission
            corresponding to the overlapping clone as we submit sequences with
            only a small overlap as described above.
            This sequence was generated from part of bacterial clone contigs of
            human chromosome X, constructed by the Sanger Centre chromosome X
            mapping group. Further information can be found at
            http://www.sanger.ac.uk/HGP/ChrX/
            This sequence has been finished according to sequence map criteria
            as follows. An attempt is made to resolve all sequencing problems,
            such as compressions and repeats, but not necessarily within known
            annotated human repeat sequence elements (e.g. Alu). Where the
            sequence is ambiguous, there is an annotation using the 'unsure'
            feature key.
            The true left end of clone 230G1 is at 1 in this sequence. The true
            right end of clone 230G1 is at 125515.
            230G1 is from the library RPCI1 constructed at the Roswell Park
            Cancer Institute by the group of Pieter de Jong.
            For further details see http://dacpac.med.buffalo.edu/
            Location/Qualifiers
                1..125515
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /chromosome="X"
                /map="p11.3"
                /clone="230G1"
                /clone_1kb="RPCI1"
            repeat_region
                1..68
                /note="AluSp repeat: matches 234. .301 of consensus;
                incomplete repeat"
            repeat_region
                525..778
                /note="AluDo repeat: matches 47. .300 of consensus;
                incomplete repeat"
            repeat_region
                1317..1586
                /note="AluSc repeat: matches 297. .2 of consensus"
            repeat_region
                1653..1907
                /note="MIR repeat: matches 1. .241 of consensus"
            repeat_region
                1917..2071
                /note="LI repeat: matches 3981. .3830 of consensus"
            repeat_region
                2127..2357
                /note="LI repeat: matches 3812. .3573 of consensus"
            repeat_region
                2369..2500
                /note="FLAM_C repeat: matches 132. .1 of consensus"
            repeat_region
                2522..3455
                /note="LI repeat: matches 3546. .2613 of consensus"
            repeat_region
                3477..3772
                /note="AluDo repeat: matches 1. .297 of consensus"
            repeat_region
                3700..4257
                /note="match: STS AL023622"
            repeat_region
                3775..4139
                /note="LI repeat: matches 2593. .2227 of consensus"
            repeat_region
                4940..5116
                /note="AluDo repeat: matches 297. .121 of consensus;
                incomplete repeat"
            repeat_region
                5129..5429
                /note="AluSg repeat: matches 300. .1 of consensus"
            repeat_region
                5430..5560
                /note="AluDo repeat: matches 135. .1 of consensus;
                incomplete repeat"
            repeat_region
                5940..6297
                /note="LI repeat: matches 1982. .1612 of consensus"
            repeat_region
                6620..6923
                /note="AluSg repeat: matches 303. .1 of consensus"
            repeat_region
                8204..8300
                /note="MER42c repeat: matches 283. .187 of consensus"
            repeat_region
                8314..8351
                /note="19 copies of 2 mer 82 % conserved"
            repeat_region
                8395..8927
                /note="MER21B repeat: matches 14. .541 of consensus"
            repeat_region
                8950..9250
                /note="AluSp repeat: matches 2. .303 of consensus"
            repeat_region
                9263..9414
                /note="MER39 repeat: matches 226. .380 of consensus"
            repeat_region
                9423..9725
                /note="AluSx repeat: matches 302. .1 of consensus"
            repeat_region
                9726..10024
                /note="AluDo repeat: matches 296. .3 of consensus"
            repeat_region
                10407..10702
                /note="AluSx repeat: matches 1. .297 of consensus"
            repeat_region
                10721..10781
                /note="LI repeat: matches 3705. .3766 of consensus"
            repeat_region
                10786..11101
                /note="AluDo repeat: matches 5. .301 of consensus"
            repeat_region
                11116..11482
                /note="LI repeat: matches 4285. .4647 of consensus"
            repeat_region
                11619..11917
                /note="AluDo repeat: matches 301. .1 of consensus"
            repeat_region
                12241..12536
                /note="AluSg repeat: matches 297. .1 of consensus"
            repeat_region
                12791..13084
                /note="AluDo repeat: matches 1. .302 of consensus"
            repeat_region
                14127..14359
                /note="MIR repeat: matches 260. .23 of consensus"
            repeat_region
                14360..14663
                /note="AluDo repeat: matches 301. .1 of consensus"
            repeat_region
                14664..14945
                /note="AluSg repeat: matches 297. .11 of consensus"
            repeat_region
                19264..19408
                /note="MIR2 repeat: matches 145. .1 of consensus"
            repeat_region
                19777..20103
                /note="MER31 repeat: matches 3. .328 of consensus"
            repeat_region
                21907..21951
                /note="MIR2 repeat: matches 145. .97 of consensus"
            repeat_region
                22064..22179
                /note="MIR repeat: matches 134. .18 of consensus"
            repeat_region
                22192..22400
                /note="MIR repeat: matches 259. .30 of consensus"
            repeat_region
                22889..23026
                /note="LIPB3 repeat: matches 746. .883 of consensus"
            repeat_region
                23027..23321
                /note="AluSx repeat: matches 298. .1 of consensus"
            repeat_region
                23329..23626
                /note="AluSg repeat: matches 298. .1 of consensus"
            repeat_region
                23653..23788
                /note="FLAM_C repeat: matches 1. .133 of consensus"
            repeat_region
                23893..24071
                /note="MIR repeat: matches 262. .90 of consensus"
            repeat_region
                24523..24575
                /note="MIR repeat: matches 152. .102 of consensus"
            repeat_region
                26634..26710
                /note="MADE1 repeat: matches 1. .80 of consensus"
            repeat_region
                27043..27125
                /note="LIMB8 repeat: matches 922. .840 of consensus"
            repeat_region
                <31042..>332667
                /note="match: multiple ESTs; match: H19527 H19455 H12759
                A114161 H20205; match: H12802 AA806420 D55951 FJ0957
                R22809; match: AA477696 H59758 AA904128 AA32774 T85326;
                match: AA687697 H48153 H20397 R83683 AA454735; match:
                AA182638 AA222007 H44069 H21221 AA331798; match: H46375
                H19843 A573937 AA868327 AA865442; match: AA570700 H44150
                H20192 T88882 AA015618; match: H20383 T15653 H49938 H59757

```

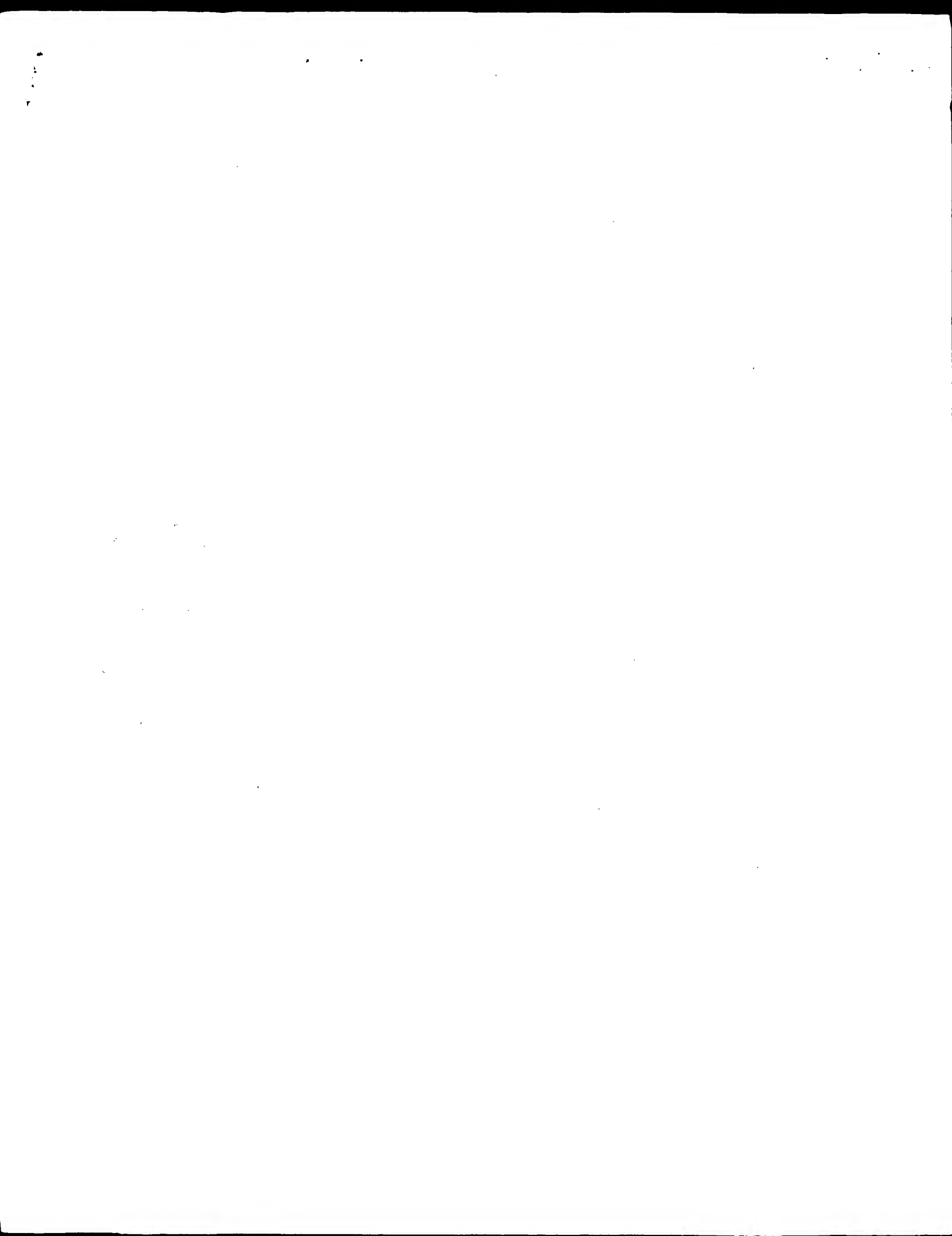
Search completed: Tue Jan 18 14:34:00 2000
Job time : 4528 secs.

```
misc_feature
  match: H51394 AA573455"
  <32049..32643
  /note="match: GSS G23742 G06335 G22442"
  32725..32762
  /note="19 copies of 2 mer 97 % conserved"
repeat_region
  <32821..32877
  /note="match: multiple ESTs; match: AA282955 AA139749
  AA571767 AA464040 AA824013; match: 221451 AA895634 H34312
  F23024 W46769 H34604; match: AA719679 W39769 AA592343
  AA92506 W54246; match: AA173069 AA425177 AA073466 M85747
  AA768214; match: AA529912 AA680646 AA569136 AA896025
  AA795810; match: AA152133 N6316 H40687 C81397 D57292;
  match: AA800504 AA273500 N34368 AA416411 AA468169; match:
  AA177766 AA634885 AA807739 C86673 N49127; match: AA175033
  HSU46225 AA798480 AA272163 M79073; match: AA219027
  AA105708 AA412536 C81242 AA003677; match: AA475009
  AA882016 H86687 C81219 H33662; match: Z98492 AA233775
  T08746 AA433377 AA619526; match: R69671 AA435188 AA760097
  H99588 AA763904"
prim_transcript
  <32961..329746
  /note="match: multiple ESTs; match: H16234 H59757 AA182732
  AA477696 R23011; match: T85154 AA573937 R83682 AA54787
  AA236603; match: H45737 T77405 F13363 N76413 W85093"
  35727..35875
  /note="MIR repeat: matches 8..167 of consensus"
  <37193..372756
  /note="match: multiple ESTs; match: R16511 T19961 R41881
  N48906 T19860; match: AA458501 H05270 AA429944; match:
  AA144162 AA119507 AA851630 AA114188 AA866788; match:
  AA104043 W88083 AA231365 AA637261 W98335; match: AA434797
  AA518513 AA119507 AA800636 AA866788; match: R17266 R96026
  AA346254 AA134624 R71984"
  37462..37572
  /note="MIR repeat: matches 146..34 of consensus"
  38310..38485
  /note="MIR repeat: matches 260..84 of consensus"
  39291..39412
  /note="MIR repeat: matches 188..68 of consensus"
  39826..40528
  /note="MER44C repeat: matches 728..4 of consensus"
  41470..41739
  /note="AluSq repeat: matches 303..35 of consensus;
  incomplete repeat"
prim_transcript
  43206..43642
  /note="match: 3' EST AA033820 clone 375753"
  44018..44331
  /note="match: 3' EST AA033820 clone 375753"
repeat_region
  44018..44331
  /note="match: 3' EST AA033820 clone 375753"

Note: remainder of annotations omitted.

Query Match      9.2%: Score 223; DB 29; Length 125515;
Best Local Similarity 86.3%; Pred. No. 9.48e-147;
Matches 258; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

Db 70117 TGGCCAGGCGTGGCTGCTGCTGTAATCCTAGACCTTTGGAGGCCAAAGCGGTGG 70176
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
QY 2144 tggccagcgcgagtgctcacacgttaatcccaacacttgggagggccaagtggtg 2203
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
Db 70177 ATCAGCTGAGTGGGAGTTCAAGACCAAGCTGGCCCAATGTGTAACCCATCTTAC 70236
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
QY 2204 atcatctaagggcgaggttcaagaccagctgagcaacggtgagaaacccatctcac 2263
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
Db 70237 TAAAAATACAAATTAGCCAGGTGTGGTGGCAATGCCGTGTAATCCAGCTACTTGGGAG 70296
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
QY 2264 tnaaaatacnaaatagccggcggtgtgtagcgatgctgtanacctggtactcagagag 2323
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
Db 70297 GCTGAGCGAGAGATTTGCTTGAACCTGGAGCGGAGGTTGCACTGAGCCAAAGATCCGG 70356
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
QY 2324 gccgagggcagaanaatttctgaactgggagggcagaggttggtgagcccgagcg 2383
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
Db 70357 CCATTGCACTCCAGCCTGGGCAACAGAACAACTCCCTCAAAAAA70415
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
QY 2384 ccattgactccagcctgggttaacagacaactctgtccaaaaa2442
```



GN HYEM-L.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
OC CARNIVORIA; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RA HARROP J.A., McDONNELL P.C., BRIGHAM-BURKE M., IYN S.D., MINTON J.,
RA TAN K.B., DEDE K., SPAMPANATO J., SILVERMAN C., HENSLEY P.,
RA DIPRINZIO R., EMERY J.G., DEEN K., EICHMAN C., CHABOT-FLETCHER M.,
RA TRUNER A., YOUNG P.R.;
RT HYEM-L, a novel ligand for HYEM/TR2, stimulates NF-kB-dependent
RT transcription, proliferation of T cells and inhibition of HT29 cell
RT growth.*
RL SUBMITTED (JUL-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RA SHINTANI Y., NISHI K., LYNN S.D., YOUNG P.R.;
RL SUBMITTED (MAY-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: AF064090; G3283356; -;
SQ SEQUENCE 240 AA; 26350 MW; 3490EB77 CRC32;

Query Match	16.4%;	Score 216;	DB 4;	Length 240;
Best Local Similarity	32.9%;	Pred. No. 1.06e-22;		
Matches	46;	Conservative	38;	Mismatches 42;
			Indels 14;	Gaps 8

[illegible]

RESULT	3	PRELIMINARY:	PR1:	240 AA.
ID	043557			
AC	043557:			
DT	01-JUN-1998	(TREMBLERL. 06, CREATED)		
DT	01-JUN-1998	(TREMBLERL. 06, LAST SEQUENCE UPDATE)		
DT	01-JUN-1998	(TREMBLERL. 06, LAST ANNOTATION UPDATE)		
DE	TUMOR NECROSIS FACTOR SUPERFAMILY MEMBER LIGHT.			
OS	HOMO SAPIENS (HUMAN).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;			
OC	CATARRHINI; HOMINIDAE; HOMO.			
OC	[1]			
RP	SEQUENCE FROM N.A.			
RA	MAURI D.N., EBERH R., MONTGOMERY R.I., KOCHER K.D., CHEUNG T.C.,			
RA	YU G.-L., RUBEN S., MURPHY M., ELSENBERG R.J., COHEN G.H., SPEAR P.G.,			
RA	WARE C.F.,			
RL	IMMUNITY 8:21-30(1998).			
DR	EMBL AF036581, G2815624, -.			
QC	SEQUENCE 240 AA; 26351 MW; 444B603A CRC32;			

Query Match	15.8%;	Score 208;	DB 4;	Length 240;
Best Local Similarity	32.1%;	Pred. No. 4.26e-21;		
Matches	45;	Conservative	38;	Mismatches 43;
			Indels 14;	Gaps 8

[illegible]

RESULT	4		
ID	077764	PRELIMINARY;	PT; 233 AA.
AC	077764;		
DT	01-NOV-1998 (TREMBLREL, 08, CREATED)		
DT	01-NOV-1998 (TREMBLREL, 08, LAST SEQUENCE UPDATE)		
DT	01-NOV-1998 (TREMBLREL, 08, LAST ANNOTATION UPDATE)		
DE	TUMOUR NECROSIS FACTOR ALPHA.		
GN	TNF.		
OS	MACROPUS EUGENII (TAMMAR WALLABY).		
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; METATHERIA;		
OC	DIPROTODONTIA; MACROPODIDAE; MACROPUS.		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=MAMMARY-ASSOCIATED LYMPH NODE;		
RA	HARRISON G.A., BROUGHTON M.J., DEANE E.M., COOPER D.W.;		
RL	SUBMITTED (MAR-1998) TO EMBL/GENBANK/DBJ DATA BANKS.		
DR	EMBL: AF055915; G3659513; -		
DR	PROSITE, PS00251; TNF.1.1.		
SO	SEQUENCE 233 AA; 25696 MW; 5B4C4E1E CRC32;		

Query Match	15.48;	Score 203;	DB 6;	Length 233;
Best Local Similarity	25.78;	Pred. No. 4.21e-20;		
Matches	35;	Conservative	39;	Mismatches 57;
			Indels	5;
			Gaps	4.

[illegible]

RESULT	5	PRELIMINARY;	PRT;	234 AA.
ID	Q28320			
AC	Q28320;			
DT	01-NOV-1996	(TREMBLREL. 01, CREATED)		
DT	01-NOV-1996	(TREMBLREL. 01, LAST SOURCE UPDATE)		
DT	01-JAN-1999	(TREMBLREL. 09, LAST ANNOTATION UPDATE)		
DE	TNF ALPHA.			
OS	CAPRA HIRCUS (GOAT).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;			
NC	ARTIODACTYLIA; RUMINANTIA; PECORA; BOVIDEA; BOVIDAE; CAPRINAE; CAPRA			
RC	[1]			
RP	SEQUENCE FROM N.A.			
RA	TAKAKURA H., MORI Y., TATSUMI M.;			
RT	"Molecular cloning of caprine TNF-alpha cDNA and its expression in			
RT	E.coli and insect cells."			
RL	SUBMITTED (JUL-1996) TO EMBL/GENBANK/DBJ DATA BANKS.			
DR	EMBL: D86587; D1013817; TO EMBL/GENBANK/DBJ DATA BANKS.			
DR	PROSITE: P500251; TNF_L; 1.			
DR	PIRAT: P00229; TNF; 1.			
Q0	SEQUENCE 234 AA; 2519 MW; C6424744 CRC32;			

Query Match	15.2%;	Score 200;	DB 6;	Length 234;
Best Local Similarity	28.7%;	Pred. No. 1.66e-19;		
Matches	39;	Conservative	31;	Mismatches 61;
			Indels	5;
			Gaps	5;

[illegible]


```

OY 42 WEHELGLAFTKRNMTTKFLLIPESGDIYFISQVTFEGMTSECEIRQAGRPKNPDSIT 101
DB 168 VVVTIKTSIKIPSSHTLMKSGSTKMSGNSEFHYFSINVGFFKLRSGEELISIEVSNPSLL 227
OY 102 VVITKVTDSYEPPTQLMGTKSV-C-EVGSNM-FQPIYLGAMFSLQEGDKLMVNVSDISLV 159
DB 228 D-PDODATYFGAF 239
OY 160 DYTEKEDKTEFGAF 172

RESULT 13
ID 070332 PRELIMINARY; PRT; 216 AA.
AC 070332;
DT 01-JAN-1998 (TREMBLREL. 07, CREATED)
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE TUMOR NECROSIS FACTOR-ALPHA (FRAGMENT).
OS MESOCRICEUTUS AURATUS (GOLDEN HAMSTER).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;
SCUROGNATHI; MURIDAE; CRICETINAE; MESOCRICEUTUS.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SPLEEN;
RA MEDLINE; 98234044.
RA MELBY P.C., TRYON V.V., CHANDRASEKAR B., FREEMAN G.L.;
RT "Cloning of Syrian hamster (Mesocricetus auratus) cytokine cDNAs and
RT analysis of cytokine mRNA expression in experimental visceral
RT leishmaniasis.";
RL INFECT. IMMUN. 66:2135-2142(1998).
DR EMBL; AF046215; G3005109; -.
DR PROSITE; PS00251; TNF_1; 1.
FT NON_TER 1
FT NON_TER 216
FT SEQUENCE 216 AA; 23793 MW; 79A08367 CRC32;
SQ

Query Match 11.6%; Score 152; DB 11; Length 216;
Best Local Similarity 27.6%; Pred. No. 2,27e-10;
Matches 34; Conservative 31; Mismatches 54; Indels 4; Gaps 4;

DB 95 LEWLSHRANALLANGMSJKNQVLPADGLVYSQVLFRRQ-G-CPSYVLLHTHSRIA 152
OY 40 LHWEHLGLAFTKRNMTTKFLLIPESGDIYFISQVTFEGMTSECEIRQAGRPKNPDS 99
DB 153 VSYE-DVNVLLSAIKSPKPEGEELK-P-WYEPYIYGVOLEKGRJLSAEVNLPRYL 210
OY 100 IIVITKVTDSYEPPTQLMGTKSVCEVGSNMFPQIYLGAMFSLQEGDKLMVNVSDISLV 159
DB 211 DPA 213
OY 160 DYT 162

RESULT 14
ID 035306 PRELIMINARY; PRT; 316 AA.
AC 035306;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE TNF-RELATED ACTIVATION-INDUCED CYTOKINE (RANKL).
GN TRANCE OR RANKL.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;
SCUROGNATHI; MURIDAE; MURINAE; MUS.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=FROM MARROW;
RA MEDLINE; 98032977.
RA ANDERSON D.M., MARASKOVSKY E., BILLINGSLEY W.L., DOUGALL W.C.,
RA TOMETSKO M.E., ROUX E.R., TEEPE M.C., DUBOSE R.F., COSMAN D.,
RA GALIBERT L.;
RT "A homologue of the TNF receptor and its ligand enhance T-cell growth
RT and dendritic-cell function.";

```

```

RL NATURE 390:175-179(1997).
DR EMBL; AF019048; G2612924; -.
DR MGD; MGI:1100089; TRANCE.
DR PFAM; PF00229; TNF; 1.
SQ SEQUENCE 316 AA; 35002 MW; 97DA4504 CRC32;

Query Match 9.9%; Score 130; DB 11; Length 316;
Best Local Similarity 26.3%; Pred. No. 1.71e-06;
Matches 35; Conservative 33; Mismatches 55; Indels 10; Gaps 9;

DB 186 WYHDRGMA-KISMTTNGCKLRVNOGFFYLLYANICFRHETS-GSV-PTD--Y-LQ-LM 238
OY 42 WEHELGLAFTKRNMTTKFLLIPESGDIYFISQVTFEGMTSECEIRQAGRPKNPDSIT 101
DB 239 VVVTIKTSIKIPSSHTLMKSGSTKMSGNSEFHYFSINVGFFKLRAGEISIOVSNPSLL 298
OY 102 VVITKVTDSYEPPTQLMGTKSV-C-EVGSNM-FQPIYLGAMFSLQEGDKLMVNVSDISLV 159
DB 299 D-PDODATYFGAF 310
OY 160 DYTEKEDKTEFGAF 172

RESULT 15
ID 035235 PRELIMINARY; PRT; 316 AA.
AC 035235;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE TNF-RELATED ACTIVATION-INDUCED CYTOKINE (TRANCE).
GN TRANCE.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;
SCUROGNATHI; MURIDAE; MURINAE; MUS.
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE; 97460112.
RA WONG B.R., RHO J., ARRON J., ROBINSON E., ORLINICK J., CHAO M.,
RA KALACHIKOV S., CAVANI E., BARTLETT F.S. III, FRANKEL W.N., LEE S.Y.,
RA CHOI Y.;
RT "TRANCE is a novel ligand of the tumor necrosis factor receptor
RT family that activates c-Jun N-terminal kinase in T cells.";
RL J. BIOL. CHEM. 272:25190-25194(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC MEDLINE; 98227661.
RA LACEY D.L., TIMMS E., TAN H.-L., KELLEY M.J., DUNSTAN C.R.,
RA BURGESS T., ELLIOTT R., COLOMERO A., ELLIOTT G., SCULLY S., HSU H.,
RA SULLIVAN J., HAWKINS N., DAVY E., CAPPAPELLI C., ELI A., QIAN Y.-X.,
RA KAUFMAN S., SAROSI I., SHALHOUB V., SENALDI G., GUO J., DELANEY J.,
RA BOYLE W.J.;
RT "Osteoprotegerin ligand is a cytokine that regulates osteoclast
RT differentiation and activation.";
RL CELL 93:165-176(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=BONE MARROW;
RX MEDLINE; 98188248.
RA YASUDA H., SHIMA N., NAKAGAWA N., YAMAGUCHI K., KINOSAKI M.,
RA MCHIZUKI S., TOMOTASU A., YANO K., GOTO M., MURAKAMI A., TSUDA E.,
RA MORINAGA T., HIGASHIO K., UDAGAWA N., TAKAHASHI N., SUDA T.;
RT "Osteoclast differentiation factor is a ligand for
RT osteoprotegerin/osteoclastogenesis-inhibitory factor and is identical
RT to TRANCE/RANKL.";
RL PROC. NATL. ACAD. SCI. U.S.A. 95:3597-3602(1998).
DR EMBL; AF053713; G3057148; -.
DR EMBL; AF013170; G2411498; -.
DR EMBL; AB008426; D1026353; -.
DR MGD; MGI:1100089; TRANCE.
DR PFAM; PF00229; TNF; 1.
SQ SEQUENCE 316 AA; 34944 MW; F76EC806 CRC32;

Query Match 9.9%; Score 130; DB 11; Length 316;

```

Best Local	Similarity	26.38;	Pred. No.	1.71e-06;					
Matches	35;	Conservative	33;	Mismatches	55;	Indels	10;	Gaps	9;

Db 186 WYHDRGWA-KISNMTLSNGKLRVNQDGFYLYLANICFRHETS-GSV-PTD--Y-LQ-LM 238

QY 42 WEHELGLAFTKNRNNYTNKFLLPESGDYFIYSQVTFRGMTSECSIRQAGRPNKPDST 101

Db 239 YVVKTSIKIPSSHNLMMGGSTKNWSGENSEPHYYSINVGGEFKLRAGEEISIQYSNPSLL 298

QY 102 VVITKVTDSYPEPTQLMGTKSVC-EVGSNW-FQPIYLGAMESLQEGDKLMVNSDISLV 159

Db 299 D-PDQDATYFGAF 310

QY 160 DYTKEKDTFFGA 172

Search completed: Tue Jan 18 13:22:22 2000
Job time : 72 secs.

 M I S E R E (TM)

Release 3.1a John F. Collins, Bioinformatics Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

MPsrch_dp protein - protein database search, using Smith-Waterman algorithm
 Run on: Tue Jan 18 13:20:29 2000; MasPar time 6.81 Seconds
 Tabular output not generated.

Title: >US-09-246-129-2
 Description: (1-174) from US09246129A.pep
 Perfect Score: 1314
 Sequence: 1 MKRFLSKYSEPMRKILFL.....DISLVDYTKEDKTFEGAFLL 174

Scoring table: PAM 150
 Gap 11

Searched: 77977 seqs, 28268293 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: swiss-prot37
 1:swissprot

Statistics: Mean 45.305; Variance 75.582; scale 0.599

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.	
1	242	18.4	202	1	TNFB_RAT	LYMPHOTOXIN-ALPHA PREC	4.89e-29
2	241	18.3	202	1	TNFB_MOUSE	LYMPHOTOXIN-ALPHA PREC	8.00e-29
3	235	17.9	291	1	TRAI_MOUSE	TNF-RELATED APOPTOSIS	1.52e-27
4	232	17.7	279	1	FASL_MOUSE	FAS ANTIGEN LIGAND	6.56e-27
5	231	17.6	281	1	TRAI_HUMAN	TNF-RELATED APOPTOSIS	1.07e-26
6	229	17.4	197	1	TNFB_HUMAN	LYMPHOTOXIN-ALPHA PREC	2.83e-26
7	229	17.4	204	1	TNFB_BOVIN	LYMPHOTOXIN-ALPHA PREC	2.83e-26
8	220	16.7	281	1	FASL_HUMAN	FAS ANTIGEN LIGAND (AP	2.20e-24
9	217	16.5	204	1	TNFB_PIG	LYMPHOTOXIN-ALPHA PREC	9.32e-24
10	212	16.1	278	1	FASL_RAT	FAS ANTIGEN LIGAND	1.02e-22
11	200	15.2	205	1	TNFB_HUMAN	LYMPHOTOXIN-ALPHA PREC	3.00e-20
12	200	15.2	233	1	TNFA_MACFA	TUMOR NECROSIS FACTOR	3.00e-20
13	198	15.1	233	1	TNFA_TRIUV	TUMOR NECROSIS FACTOR	7.68e-20
14	197	15.0	233	1	TNFA_MACMU	TUMOR NECROSIS FACTOR	1.23e-19
15	197	15.0	233	1	TNFA_PAPHU	TUMOR NECROSIS FACTOR	1.23e-19
16	196	14.9	193	1	TNFA_CAPHI	TUMOR NECROSIS FACTOR	1.96e-19
17	196	14.9	232	1	TNFA_PIG	TUMOR NECROSIS FACTOR	1.96e-19
18	194	14.8	233	1	TNFA_HUMAN	TUMOR NECROSIS FACTOR	4.98e-19
19	191	14.5	233	1	TNFA_PAPSP	TUMOR NECROSIS FACTOR	2.01e-18
20	191	14.5	233	1	TNFA_CANPA	TUMOR NECROSIS FACTOR	2.01e-18
21	191	14.5	306	1	TNFC_MOUSE	LYMPHOTOXIN-BETA (LT-B	2.01e-18
22	189	14.4	234	1	TNFA_SHEEP	TUMOR NECROSIS FACTOR	5.07e-18
23	189	14.4	235	1	TNFA_MOUSE	TUMOR NECROSIS FACTOR	5.07e-18

RESULT ID	1	TNFB_RAT	STANDARD	PRT	202 AA	ALIGNMENTS
AC	006332					
DT	01-JUN-1994	(REL. 29, CREATED)				
DT	01-JUN-1994	(REL. 29, LAST SEQUENCE UPDATE)				
DT	01-NOV-1997	(REL. 35, LAST ANNOTATION UPDATE)				
DE	LYMPHOTOXIN-ALPHA PRECURSOR (LT-ALPHA) (TNF-BETA).					
GN	LTA OR TNFB.					
OS	RATTUS NORVEGICUS (RAT).					
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;					
OC	RODENTIA; SCIUROGNATHI; MORIDAE; MORINAE; RATTUS.					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RC	STRAIN-SPRAGUE-DAWLEY; TISSUE-TESTIS;					
RX	MEDLINE: 94040766.					
RA	KWON J., CHUNG I.Y., BENVENISTE E.N.;					
RT	"Cloning and sequence analysis of the rat tumor necrosis					
RT	factor-encoding genes.";					
RL	GENE 132:227-236(1993).					
CC	- FUNCTION: LYMPHOTOXIN IS A CYTOKINE PRODUCED BY LYMPHOCYTES WHICH					
CC	IS CYTOTOXIC FOR A WIDE RANGE OF TUMOR CELLS IN VITRO AND IN VIVO.					
CC	- SUBUNIT: HETEROTRIMER OF EITHER TWO BETA AND ONE ALPHA SUBUNITS					
CC	OR (LESS PREVALENT) TWO ALPHA AND ONE BETA SUBUNITS.					
CC	- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.					
CC	-----					
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration					
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -					
CC	the European Bioinformatics Institute. There are no restrictions on its					
CC	use by non-profit institutions as long as its content is in no way					
CC	modified and this statement is not removed. Usage by and for commercial					
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/					
CC	or send an email to license@isb-sib.ch).					
CC	-----					
DR	EMBL: L00981; G205255; -					
DR	PIR: JN0869; JN0869.					
DR	PROSITE: PS00251; TNF_1; 1.					
DR	PROSITE: PS50049; TNF_2; 1.					
DR	PRAM; PF00229; TNF; 1.					
KW	CYTOKINE; GLYCOPROTEIN; CYTOTOXIN; SIGNAL.					
FT	SIGNAL	1	33			
FT	CHAIN	34	202			
FT	CARBOHYD	94	94			
FT	SEQUENCE	202 AA;	2211 MM;	6E2C05C4 CRG32;		

Query Match 18.4%; Score 242; DB 1; Length 202;
 Best Local Similarity 36.5%; Pred. No. 4.89e-29;

RA WILEY S.R., SCHOOLEY K., SMOLAK P.J., DIN W.S., HUANG C.-P.,
 RA NICHOL J.K., SUTHERLAND G.R., DAVIS-SMITH T., RADUCH C., SMITH C.A.,
 RA GOODWIN R.G.;
 RT "Identification and characterization of a new member of the TNF
 RT family that induces apoptosis.";
 RL IMMUNITY 3:673-682(1995).
 CC -1- FUNCTION: INDUCES APOPTOSIS.
 CC -1- SUBUNIT: HOMOTRIMER (POTENTIAL).
 CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN (POTENTIAL).
 CC -1- TISSUE SPECIFICITY: WIDESPREAD.
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U37522; G1149560; -.
 DR MGD: MGI:107414; TRAIL.
 DR PROSITE: PS00251; TNF_1; 1.
 DR PROSITE: PS50049; TNF_2; 1.
 DR PRAM: PF00229; TNF; 1.
 DR CYTOKINE: TRANSMEMBRANE; SIGNAL-ANCHOR; APOPTOSIS.
 KW DOMAIN 1 17 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 18 38 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
 FT DOMAIN 39 291 EXTRACELLULAR (POTENTIAL).
 FT CARBOHYD 52 52 POTENTIAL.
 SQ SEQUENCE 291 AA: 33477 MW: F9906FBE CRC32;
 Query Match 17.9%; Score 235; DB 1; Length 291;
 Best Local Similarity 31.4%; Pred. No. 1,52e-27;
 Matches 43; Conservative 35; Mismatches 51; Indels 8; Gaps 8;
 Db 158 WESSRKGHSL-NHVLFRNCGLYIEDEGLYIYSQYFFQEAR-DASWMSKDYRKQ 215
 QY 42 WEHEL-GLAFTKRMNYTNKFLIPESGYFIYSQYTFRGMSECEIRQAGRPN-KPDS 99
 Db 216 LVQYIKYKT-SYDPPIVLMKSARNSQMSRDAEYGLYSIQGGLFELKKNDRIFVSTNH 274
 QY 100 ITVTITKVDSTPEPTQLMGTSVC-EVGSNM-FQPIILGAFSLQESDKLMVNSDIS 157
 Db 275 LMDLQEA-SFGGAPLI 290
 QY 158 LVDTYKEDKFFGAPLL 174
 RESULT 4
 ID FASL_MOUSE STANDARD; PRT; 279 AA.
 AC P41047;
 DT 01-FEB-1995 (REL. 31, CREATED)
 DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
 DE FAS ANTIGEN LIGAND.
 GN APTILGI OR FASL OR GLD.
 OS MUS MUSCULUS (MOUSE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA.
 OC RODENTIA; SCIROGNATHI; MURIDAE; MURINAE; MUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 94185175.
 RA TAKAHASHI T., TANAKA M., BRANNAN C.I., JENKINS N.A., COPELAND N.G.,
 RA SUDA T., NAGATA S.;
 RT "Generalized lymphoproliferative disease in mice, caused by a point
 RT mutation in the fas ligand.";
 RL CELL 76:969-976(1994).
 CC [2]
 RN SEQUENCE FROM N.A., AND 3D-STRUCTURE MODELLING.
 RC STRAIN=C57BL/6;
 RX MEDLINE: 95388076
 RA PEITSCH M.J., TSCHOPP J.J.;

RT "Comparative molecular modelling of the Fas-ligand and other members
 RT of the TNF family.";
 RL MOL. IMMUNOL. 32:761-772(1995).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 95196085.
 RA LYNCH D.H., WATSON M.L., ALDERSON M.R., BAUM P.R., MILLER R.E.,
 RA TOUGH T., GIBSON M., DAVIS-SMITH T., SMITH C.A., HUNTER K.;
 RT "The mouse Fas-ligand gene is mutated in gld mice and is part of a
 RT TNF family gene cluster.";
 RL IMMUNITY 1:131-136(1994).
 RN [4]
 RP CHARACTERIZATION OF VARIANT GLD.
 RX MEDLINE: 96091792.
 RA HANE M., PEITSCH M.C., IMLER M., SCHROETER M., LOWIN B.,
 RA ROUSSEAU M., BRON C., RENNO T., FRENCH L., TSCHOPP J.;
 RT "Characterization of the non-functional Fas ligand of gld mice.";
 RL INT. IMMUNOL. 7:1381-1386(1995).
 CC -1- FUNCTION: CYTOKINE THAT BINDS TO FAS ANTIGEN, A RECEPTOR THAT
 CC TRANSDUCES THE APOPTOTIC SIGNAL INTO CELLS. MAY BE INVOLVED IN
 CC CYTOTOXIC T CELL MEDIATED APOPTOSIS AND IN T CELL DEVELOPMENT.
 CC FAS-ANTIGEN MEDIATED APOPTOSIS MAY HAVE A ROLE IN THE INDUCTION OF
 CC PERIPHERAL TOLERANCE, IN THE ANTIGEN-STIMULATED SUICIDE OF MATURE
 CC T CELLS, OR BOTH.
 CC -1- SUBUNIT: HOMOTRIMER (PROBABLE).
 CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MAY BE RELEASED
 CC INTO THE EXTRACELLULAR FLUID, PROBABLY BY CLEAVAGE FORM THE CELL
 CC SURFACE.
 CC -1- DISEASE: A DEFICIENCY IN THIS PROTEIN IS THE CAUSE OF GENERALIZED
 CC LYMPHOPROLIFERATION DISEASE (GLD), AN AUTOSOMAL RECESSIVE DISEASE
 CC RESPONSIBLE FOR LYMPHADENOPATHY AND AUTOANTIBODY PRODUCTION.
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U06948; G473565; -.
 DR EMBL: U10984; G511222; -.
 DR MGD: S76752; G913760; -.
 DR MGD: MGI:99255; FASL.
 DR PROSITE: PS00251; TNF_1; 1.
 DR PROSITE: PS50049; TNF_2; 1.
 DR PRAM: PF00229; TNF; 1.
 DR HSSP: P01375; TNF.
 DR CYTOKINE: TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR; APOPTOSIS;
 KW DISEASE MUTATION.
 FT DOMAIN 1 78
 FT TRANSMEM 79 100 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 101 279 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
 FT DOMAIN 4 69 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 45 51 PRO-RICH.
 FT DISULFID 200 231 POLY-PRO.
 FT VARIANT 273 273 BY SIMILARITY.
 FT CARBOHYD 117 117 F->L (TN GLD; ABOLISH BINDING OF FASL TO
 FT CARBOHYD 182 182 ITS RECEPTOR).
 FT CARBOHYD 248 248 POTENTIAL.
 FT CARBOHYD 258 258 POTENTIAL.
 SQ SEQUENCE 279 AA: AFD64DAB CRC32;
 Query Match 17.7%; Score 232; DB 1; Length 279;
 Best Local Similarity 35.3%; Pred. No. 6,56e-27;
 Matches 54; Conservative 32; Mismatches 50; Indels 17; Gaps 13;
 Db 143 SVALTGNPSRS-IP-LEMDTYGTALISG-VKYYKGLVNETGLYFYSAVYRGQ- 198
 QY 23 PVRKOTPTQHFKNQFPALHWEHGLAFTKRMNYTNKFLIPESGYFIYSQYTFRGM 82


```

Db 182 HL-LLSPSSVFGAFAL 197
   158 LVDYTKEDKTEFGAFLL 174

RESULT 7
ID TNER_BOVIN STANDARD: PRT: 204 AA.
AC 006600:
DT 01-JUN-1994 (REL. 29, CREATED)
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE LYMPHOTOXIN-ALPHA PRECURSOR (LT-ALPHA) (TNF-BETA).
GN LTA OR TNF.
OS BOS TAURUS (BOVINE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC ARTIODACTYLIA; RUMINANTIA; PECORA; BOVIDEA; BOVINAE; BOS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 94083525.
RA CLUDTS I., CLEUTER Y., KETTMANN R., BURNY A., DROOGMANS L.;
RT "Cloning and characterization of the tandemly arranged bovine
RT lymphotoxin and tumour necrosis factor-alpha genes.";
RL CYTOKINE 5:336-341(1993).
CC -1- FUNCTION: LYMPHOTOXIN IS A CYTOKINE PRODUCED BY LYMPHOCYTES WHICH
CC IS CYTOTOXIC FOR A WIDE RANGE OF TUMOR CELLS IN VITRO AND IN VIVO.
CC -1- SUBUNIT: HETERODIMER OF EITHER TWO BETA AND ONE ALPHA SUBUNITS
CC OR (LESS PREVALENT) TWO ALPHA AND ONE BETA SUBUNITS.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: 214137; G7971.
DR PIR: S24641; S24641.
DR PROSITE: PS00251; TNF_1; 1.
DR PROSITE: PS50049; TNF_2; 1.
DR PFAM: PF00229; TNF; 1.
DR CYTOKINE; GLYCOPROTEIN; CYTOTOXIN; SIGNAL.
FT SIGNAL 1 33 BY SIMILARITY.
FT CHAIN 34 204 LYMPHOTOXIN-ALPHA.
FT CARBOHYD 95 95 POTENTIAL.
SQ SEQUENCE 204 AA; 22098 MW; E0B38524 CRC32;

Query Match 17.4%; Score 229; DB 1; Length 204;
Best Local Similarity 35.0%; Pred. No. 2,83e-26;
Matches 48; Conservative 23; Mismatches 58; Indels 8; Gaps 8;

Db 75 STRRANDBRALRKGFSISNNSLLVPTSGLYFYVSYVFSGR-G-GFP-R-AT-PT-PL 128
   39 ALAHWEHELGLATKRNMTNFKLLIPESGDYFYISQVTFRKMSECEIRQAGRPKPD 98
Db 129 VLAHEVOLFSPOYFPHVPLLSAOKSVCGPOGPMWRYVYOGAVELLTGGDLSHTDGIS 188
   99 SITVITKVTDSYPTQLIMGTKSVCEVG-SNMFPILYLGAMSLQGDGLMNVSDIS 157
Db 189 HL-LLSPSSVFGAFAL 204
   158 LVDYTKEDKTEFGAFLL 174

RESULT 8
ID FASL_HUMAN STANDARD: PRT: 281 AA.
AC P48023:
DT 01-FEB-1996 (REL. 33, CREATED)
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)

```

```

DE FAS ANTIGEN LIGAND (APOPTOSIS ANTIGEN LIGAND) (APTL).
GN AFLIG1 OR FASL.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 95105731.
RA ALDERSON M.;
RT "Fas ligand mediates activation-induced cell death in human T
RT lymphocytes.";
RL J. EXP. MED. 181:71-77(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE: 95127560.
RA TAKAHASHI T., TANAKA M., INAZAWA J., ABE T., SUDA T., NAGATA S.;
RT "Human Fas ligand: gene structure, chromosomal location and species
RT specificity.";
RL INT. IMMUNOL. 6:1567-1574(1994).
RN [3]
RP SEQUENCE FROM N.A.
RA SCHAEFFELZLEIN C.E., POEHLMANN R., PHILIPPSEN P., EIBEL H.;
RL SUBMITTED (JUN-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE: 95071350.
RA MITA E., HAYASHI N., IIO S., TAKEHARA T., HIJIOKA T., KASAHARA A.,
RA FUSAMOTO H., KAMADA T.;
RT "Role of Fas ligand in apoptosis induced by hepatitis C virus
RT infection.";
RL BIOCHEM. BIOPHYS. RES. COMMUN. 204:468-474(1994).
RN [5]
RP SEQUENCE FROM N.A.
RA WILKINSON J.;
RL SUBMITTED (MAY-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- FUNCTION: CYTOKINE THAT BINDS TO FAS ANTIGEN, A RECEPTOR THAT
CC TRANSDUCES THE APOPTOTIC SIGNAL INTO CELLS. MAY BE INVOLVED IN
CC CYTOTOXIC T CELL MEDIATED APOPTOSIS AND IN T CELL DEVELOPMENT.
CC FAS-ANTIGEN LOCATION: TYPE II MEMBRANE PROTEIN. MAY BE RELEASED
CC PERIPHERAL TOLERANCE, IN THE ANTIGEN-STIMULATED SUICIDE OF MATURE
CC T CELLS, OR BOTH.
CC -1- SUBUNIT: HOMOTRIMER (PROBABLE).
CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MAY BE RELEASED
CC INTO THE EXTRACELLULAR FLUID, PROBABLY BY CLEAVAGE FORM THE CELL
CC SURFACE.
CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X89102; G887456; -.
DR EMBL: U08137; G624628; -.
DR EMBL: U11821; G595431; -.
DR EMBL: D38122; G1369902; -.
DR EMBL: Z96050; E320286; -.
DR MIM: 134638; -.
DR PROSITE: PS00251; TNF_1; 1.
DR PROSITE: PS50049; TNF_2; 1.
DR PFAM: PF00229; TNF; 1.
DR HSSP: P01375; 2TUN.
KM CYTOKINE; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR; APOPTOSIS;
KW DISEASE MUTATION.
FT DOMAIN 1 80 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 81 102 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
FT DOMAIN 103 281 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 4 70 PRO-RICH.
FT DOMAIN 45 65 POLY-PRO.
FT DISULFID 202 233 BY SIMILARITY.

```

```

FT CARBOHYD 184 184 POTENTIAL.
FT CARBOHYD 250 250 POTENTIAL.
FT CARBOHYD 260 260 POTENTIAL.
SQ SEQUENCE 281 AA: 31485 MW: 714BA90 CRC32:

Query Match
Best Local Similarity 35.8%; Score 220; DB 1; Length 281;
Matches 49; Conservative 30; Mismatches 43; Indels 15; Gaps 9;

Db 159 PLEMETVGVLLSG-VKKKGGVYVNGGLFVYSKYVGRQ-S-CNNLPLS--HK-- 210
OY 39 ALHWEHELGLATKNNMNTNKFLLIPESGDYFYISQVTFRMTSECSERQAGRPKNPD 98
Db 211 ---VVM-R-NSKYPQDLYMVEGKMSYCTTGQMARSSYLGAVNLTSDHLVYVSELS 265
OY 99 SITVITKVTDSYPTPTQLMGTK-SYCVGSNNFQPIYLGAMFSLQSDKLMVNSDLS 157
Db 266 LVNF-ESQTFEGLYKL 281
OY 158 LVDTKEDKTFEGAFLL 174

RESULT 9
ID TNFB_PIG STANDARD: PRT; 204 AA.
AC P26445;
DT 01-AUG-1992 (REL. 23, CREATED)
DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE LYMPHOTOXIN-ALPHA PRECURSOR (LT-ALPHA) (TNF-BETA).
GN LTA OR TNFB.
OS SCS SCROFA (PIG).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC ARTIODACTYLA; SUIFORMES; SUINA; SUIDAE; SUS.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE: 91340150.
RA KUHNERT P., WETTRICH C., PETERHANS E., PAULI U.;
RT "The porcine tumor necrosis factor-encoding genes: sequence and
RT comparative analysis.";
RL GENE 102:111-178(1991).

CC -1- FUNCTION: LYMPHOTOXIN IS A CYTOKINE PRODUCED BY LYMPHOCYTES WHICH
CC IS CYTOTOXIC FOR A WIDE RANGE OF TUMOR CELLS IN VITRO AND IN VIVO.
CC -1- SUBUNIT: HETERODIMER OF EITHER TWO BETA AND ONE ALPHA SUBUNITS
CC OR (LESS PREVALENT) TWO ALPHA AND ONE BETA SUBUNITS.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: X54859; G2133; -.
CC PIR: S17289; S17289.
CC PROSITE: PS00251; TNF_1; 1.
CC PROSITE: PS50049; TNF_2; 1.
CC PFAM: PF00229; TNF; 1.
CC CYTOKINE; GLYCOPROTEIN; CYTOTOXIN; SIGNAL.
FT SIGNAL 1 33
FT CHAIN 34 204 LYMPHOTOXIN-ALPHA.
FT CARBOHYD 95 95 POTENTIAL.
SQ SEQUENCE 204 AA: 21960 MW: 1127871C CRC32:

Query Match
Best Local Similarity 32.8%; Score 217; DB 1; Length 204;
Matches 45; Conservative 25; Mismatches 59; Indels 8; Gaps 7;

Db 75 SLWRANTRAFLRHGFLLSNNLSLVPTSGLFVYSQVFSG-EG-C--FPKAT-PT-PL 128

```

```

OY 39 ALHWEHELGLATKNNMNTNKFLLIPESGDYFYISQVTFRMTSECSERQAGRPKNPD 98
Db 129 YLAHEVQLFSSQPFVHPLLSAOKSYCPQPGWVSVYQGVFLITQDQSLTHDGP 188
OY 99 SITVITKVTDSYPTPTQLMGTKSYCVGSNNFQPIYLGAMFSLQSDKLMVNSDLS 157
Db 189 HL-LSPSSVFEFAPAL 204
OY 158 LVDTKEDKTFEGAFLL 174

RESULT 10
ID FASL-RAT STANDARD: PRT; 278 AA.
AC P36940;
DT 01-JUN-1994 (REL. 29, CREATED)
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE FAS ANTIGEN LIGAND.
GN APTLIG1 OR FASL.
OS RATIUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATIUS.
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE: 94084792.
RA SUDA T., TAKAHASHI T., GOLSTEIN P., NAGATA S.;
RT "Molecular cloning and expression of the Fas ligand, a novel member
RT of the tumor necrosis factor family.";
RL CELL 75:1169-1178(1993).

CC -1- FUNCTION: CYTOKINE THAT BINDS TO FAS ANTIGEN, A RECEPTOR THAT
CC TRANSDUCES THE APOPTOTIC SIGNAL INTO CELLS. MAY BE INVOLVED IN
CC CYTOTOXIC T CELL MEDIATED APOPTOSIS AND IN T CELL DEVELOPMENT.
CC FAS-ANTIGEN MEDIATED APOPTOSIS MAY HAVE A ROLE IN THE INDUCTION OF
CC PERIPHERAL TOLERANCE. IN THE ANTIGEN-STIMULATED SUICIDE OF MATURE
CC T CELLS, OR BOTH.
CC -1- SUBUNIT: HOMODIMER (PROBABLE).
CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MAY BE RELEASED
CC INTO THE EXTRACELLULAR FLUID, PROBABLY BY CLEAVAGE FORM THE CELL
CC SURFACE.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN ACTIVATED SPLENOCYTES AND
CC THYMOCYTES. MODERATE OR WEAK EXPRESSION FOUND IN SMALL INTESTINES,
CC KIDNEY AND LUNG.
CC -1- INDUCTION: BY PM/IONOMYCIN AND CONCAVALIN/INTERLEUKIN-2.
CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: U03470; G440179; -.
CC PROSITE: PS00251; TNF_1; 1.
CC PROSITE: PS50049; TNF_2; 1.
CC PFAM: PF00229; TNF; 1.
CC CYTOKINE; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR; APOPTOSIS.
FT DOMAIN 1 77
FT TRANSMEM 78 99
FT DOMAIN 100 278
FT DOMAIN 4 69
FT DOMAIN 45 58
FT DISULFID 199 230
FT CARBOHYD 116 116 POTENTIAL.
FT CARBOHYD 247 247 POTENTIAL.
FT CARBOHYD 257 257 POTENTIAL.
FT CARBOHYD 247 247 POTENTIAL.
FT CARBOHYD 247 247 POTENTIAL.
FT CARBOHYD 257 257 POTENTIAL.
SQ SEQUENCE 278 AA: 31140 MW: 6DC17725 CRC32:

Query Match
Best Local Similarity 16.1%; Score 212; DB 1; Length 278;

```


QY 99 STTVVITKVDTSYPPPTDLMGKTSVC-EVGSNMFQPIYLGAAMFSLQEGDKLWNVSDIS 157
 DB 190 HLVLIS-PTSTVFAGAFAL 205
 QY 158 LVDYTKEDKTEFFGAFLL 174

RESULT 12
 ID TNFA_MCFCA STANDARD: PRT: 233 AA.
 AC P793374;
 DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
 DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
 DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
 DE TUMOR NECROSIS FACTOR PRECURSOR (TNF-ALPHA) (CACHECTIN).
 GN TNFA.
 OS MACACA FASCICULARIS (CRAB EATING MACROVE) (CYNOMOLGUS MONKEY).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
 OC PRIMATES; CATARRHINI; CERCOPITHECIDAE; CERCOPITHECINAE; MACACA.
 RN [1]
 RN SEQUENCE FROM N.A.
 RA TISSUE-LYMPHOCYTES;
 RA TATSUMI M.;
 RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC -1- FUNCTION: TNF IS MAINLY SECRETED BY MACROPHAGES, IT IS A CYTOKINE
 WITH A WIDE VARIETY OF FUNCTIONS: IT CAN CAUSE CYTOLYSIS OF
 CERTAIN TUMOR CELL LINES, IT IS IMPLICATED IN THE INDUCTION OF
 CACHEXIA, IT IS A POTENT PYROGEN CAUSING FEVER BY DIRECT ACTION
 OR BY STIMULATION OF INTERLEUKIN 1 SECRETION, IT CAN STIMULATE
 CELL PROLIFERATION AND INDUCE CELL DIFFERENTIATION UNDER CERTAIN
 CONDITIONS. (BY SIMILARITY).
 CC -1- SUBUNIT: HOMOTRIMER.
 CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN
 EXTRACELLULAR SOLUBLE FORM.
 CC -1- PTM: THE SOLUBLE FORM DERIVES FROM THE MEMBRANE FORM BY
 PROTEOLYTIC PROCESSING.
 CC -1- DISEASE: CACHEXIA ACCOMPANIES A VARIETY OF DISEASES, INCLUDING
 CANCER AND INFECTION, AND IS CHARACTERIZED BY GENERAL ILL HEALTH
 AND MALNUTRITION.
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AB000513; G1794149; -
 DR PROSITE: PS00251; TNF_1; 1.
 DR PROSITE: PS50049; TNF_2; 1.
 DR PFAM: PF00229; TNF; 1.
 DR HSSP: P01375; 1A8M.
 KW CYTOKINE; CYTOTOXIN; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR.
 FT PROPEP 1 76
 FT CHAIN 77 233
 FT TRANSMEM 36 56
 FT DISULFID 145 177
 FT SEQUENCE 233 AA; 25558 MW; 856FC354 CRC32;
 SO
 Query Match 15.2%; Score 200; DB 1; Length 233;
 Best Local Similarity 28.7%; Pred. No. 3,00e-20;
 Matches 39; Conservative 34; Mismatches 58; Indels 5; Gaps 5;
 DB 102 LQMLNRAVALVANGVELDNOVLVPSBSLYLYSOVLFEKQ-G-CPS-NHVLTLHTISR 158
 QY 40 LHMHEHLGLAFTKRNRYNFKLLIPESGDYFIYSQVFRGMTSCSSEIRAGRNKPD 99
 DB 159 IIVSYQTKVLLANIRSSCRKAEDDGE-PSAMDEPYLYAGVQLTDEKLVVDTNYDEN 217
 QY 100 IIVYI-TKVTDSYPTQLMLGKTSVCVGSWMFQPIYLGAAMFSLQEGDKLWNVSDISL 158
 DB 218 LDFAEFGQLYFGAIAL 233

QY 159 VDYTKEDKTEFFGAFLL 174

RESULT 13
 ID TNFA_TRIUV STANDARD: PRT: 233 AA.
 AC P79374;
 DT 15-JUL-1998 (REL. 36, CREATED)
 DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
 DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
 DE TUMOR NECROSIS FACTOR PRECURSOR (TNF-ALPHA) (CACHECTIN).
 GN TNFA.
 OS TRICHOSURUS VULPECULA (BRUSH-TAILED POSSUM).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; METATHERIA;
 OC DIPROTODONTIA; PHALANGERIDAE; TRICHOSURUS.
 RN [1]
 RN SEQUENCE FROM N.A.
 RA MEDLINE: 96315690.
 RA WEDLOCK D.N., ALDWELL F.E., BUDDLE B.M.;
 RT "Molecular cloning and characterization of tumor necrosis factor
 alpha (TNF-alpha) from the Australian common brushtail possum,
 Trichosurus vulpecula.";
 RT Trichosurus vulpecula.";
 RL IMMUNOL. CELL. BIOL. 74:151-158(1996).
 CC -1- FUNCTION: TNF IS MAINLY SECRETED BY MACROPHAGES, IT IS A CYTOKINE
 WITH A WIDE VARIETY OF FUNCTIONS: IT CAN CAUSE CYTOLYSIS OF
 CERTAIN TUMOR CELL LINES, IT IS IMPLICATED IN THE INDUCTION OF
 CACHEXIA, IT IS A POTENT PYROGEN CAUSING FEVER BY DIRECT ACTION
 OR BY STIMULATION OF INTERLEUKIN 1 SECRETION, IT CAN STIMULATE
 CELL PROLIFERATION AND INDUCE CELL DIFFERENTIATION UNDER CERTAIN
 CONDITIONS. (BY SIMILARITY).
 CC -1- SUBUNIT: HOMOTRIMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN
 EXTRACELLULAR SOLUBLE FORM (BY SIMILARITY).
 CC -1- PTM: THE SOLUBLE FORM DERIVES FROM THE MEMBRANE FORM BY
 PROTEOLYTIC PROCESSING (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: S83283; G1881812; -
 DR EMBL: AF016102; G2425046; -
 DR PROSITE: PS00251; TNF_1; 1.
 DR PROSITE: PS50049; TNF_2; 1.
 DR PFAM: PF00229; TNF; 1.
 DR HSSP: P01375; 1A8M.
 KW CYTOKINE; CYTOTOXIN; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR.
 FT PROPEP 1 77
 FT CHAIN 78 233
 FT TRANSMEM 36 56
 FT DISULFID 145 177
 FT SEQUENCE 233 AA; 25704 MW; 873F8626 CRC32;
 SO
 Query Match 15.1%; Score 198; DB 1; Length 233;
 Best Local Similarity 27.9%; Pred. No. 7,68e-20;
 Matches 38; Conservative 40; Mismatches 33; Indels 5; Gaps 4;
 DB 102 LHWGVANTLLANGMELDNOVLVPSGLVLYVSQLYKG--SOGRESLY-LTHKISR 158
 QY 40 LHMHEHLGLAFTKRNRYNFKLLIPESGDYFIYSQVFRGMTSCSSEIRAGRNKPD 99
 DB 159 FTLISQKVTLLANIRSSCRKAEDDGE-PSAMDEPYLYAGVQLTDEKLVVDTNYDEN 217
 QY 100 IIVYI-TKVTDSYPTQLMLGKTSVCVGSWMFQPIYLGAAMFSLQEGDKLWNVSDISL 158
 DB 218 LDFAEFGQLYFGAIAL 233
 QY 159 VDYTKEDKTEFFGAFLL 174

```

RESULT 14
ID TNFA_MACMU STANDARD: PRT: 233 AA.
AC P48094:
DT 01-FEB-1996 (REL. 33, CREATED)
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE TUMOR NECROSIS FACTOR PRECURSOR (TNF-ALPHA) (CACHECTIN).
GN TNFA.
OS MACACA MULATTA (RHEUS MACAQUE).
OC EURKARYOTA: METAZOA: CHORDATA: VERTEBRATA: MAMMALIA: EUTHERIA:
OC PRIMATES: CATARRHINI: CERCOPITHECIDAE: CERCOPITHECINAE: MACACA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 96003435.
RA VILLINGER F.J., BRAR S.S., MAYNE A.E., CHIKKALA N., ANSARI A.A.;
RT "Comparative sequence analysis of cytokine genes from human and
RT nonhuman primates."
RL J. IMMUNOL. 155:3946-3954(1995).
CC -1- FUNCTION: TNF IS MAINLY SECRETED BY MACROPHAGES, IT IS A CYTOKINE
CC WITH A WIDE VARIETY OF FUNCTIONS: IT CAN CAUSE CYTOLYSIS OF
CC CERTAIN TUMOR CELL LINES, IT IS IMPLICATED IN THE INDUCTION OF
CC CACHEXIA, IT IS A POTENT PYROGEN CAUSING FEVER BY DIRECT ACTION
CC OR BY STIMULATION OF INTERLEUKIN 1 SECRETION, IT CAN STIMULATE
CC CELL PROLIFERATION AND INDUCE CELL DIFFERENTIATION UNDER CERTAIN
CC CONDITIONS.
CC -1- SUBUNIT: HOMOTRIMER.
CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN, ALSO EXISTS AS AN
CC EXTRACELLULAR SOLUBLE FORM.
CC -1- PTM: THE SOLUBLE FORM DERIVES FROM THE MEMBRANE FORM BY
CC PROTEOLYTIC PROCESSING.
CC -1- DISEASE: CACHEXIA ACCOMPANIES A VARIETY OF DISEASES, INCLUDING
CC CANCER AND INFECTION, AND IS CHARACTERIZED BY GENERAL ILL HEALTH
CC AND MALNUTRITION.
CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U19850; G644818; .
DR PROSITE: PS00251; TNF_1; 1.
DR PROSITE: PS50049; TNF_2; 1.
DR PIRAM: P00229; TNF; 1.
DR HSSP: P01375; TNF.
KW CYTOKINE; CYTOTOXIN; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR.
FT PROPEP 1 76
FT CHAIN 77 233
FT TRANSMEM 36 56
FT DISULFID 145 177
FT DISULFID 145 177
SQ SEQUENCE 233 AA; 25630 MW; 146A1B48 CRC32;

Query Match 15.0%; Score 197; DB 1; Length 233;
Best Local Similarity 28.7%; Pred. No. 1,23e-19;
Matches 39; Conservative 34; Mismatches 58; Indels 5; Gaps 5;

```

```

RESULT 15
ID TNFA_PAPHU STANDARD: PRT: 233 AA.
AC 077510;
DT 15-DEC-1998 (REL. 37, CREATED)
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE TUMOR NECROSIS FACTOR PRECURSOR (TNF-ALPHA) (CACHECTIN).
GN TNFA.
OS PAPIO HAMADRYAS URSINUS (CHACMA BABOON).
OC EURKARYOTA: METAZOA: CHORDATA: VERTEBRATA: MAMMALIA: EUTHERIA:
OC PRIMATES: CATARRHINI: CERCOPITHECIDAE: CERCOPITHECINAE: PAPIO.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 98147379.
RA HAUSER S.B., REDL H., SCHLAG G., GIROIR B.P.;
RT "Complementary DNA (cDNA) sequence of baboon tumor necrosis factor
RT alpha."
RL MOL. IMMUNOL. 34:1041-1042(1997).
CC -1- FUNCTION: TNF IS MAINLY SECRETED BY MACROPHAGES, IT IS A CYTOKINE
CC WITH A WIDE VARIETY OF FUNCTIONS: IT CAN CAUSE CYTOLYSIS OF
CC CERTAIN TUMOR CELL LINES, IT IS IMPLICATED IN THE INDUCTION OF
CC CACHEXIA, IT IS A POTENT PYROGEN CAUSING FEVER BY DIRECT ACTION
CC OR BY STIMULATION OF INTERLEUKIN 1 SECRETION, IT CAN STIMULATE
CC CELL PROLIFERATION AND INDUCE CELL DIFFERENTIATION UNDER CERTAIN
CC CONDITIONS.
CC -1- SUBUNIT: HOMOTRIMER.
CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN, ALSO EXISTS AS AN
CC EXTRACELLULAR SOLUBLE FORM.
CC -1- PTM: THE SOLUBLE FORM DERIVES FROM THE MEMBRANE FORM BY
CC PROTEOLYTIC PROCESSING.
CC -1- DISEASE: CACHEXIA ACCOMPANIES A VARIETY OF DISEASES, INCLUDING
CC CANCER AND INFECTION, AND IS CHARACTERIZED BY GENERAL ILL HEALTH
CC AND MALNUTRITION.
CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF019963; G3417555; .
DR PROSITE: PS00251; TNF_1; 1.
DR PROSITE: PS50049; TNF_2; 1.
DR CYTOKINE; CYTOTOXIN; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR.
KW PROPEP 1 76
FT CHAIN 77 233
FT TRANSMEM 36 56
FT DISULFID 145 177
FT DISULFID 145 177
SQ SEQUENCE 233 AA; 25658 MW; E38CF6E8 CRC32;

Query Match 15.0%; Score 197; DB 1; Length 233;
Best Local Similarity 28.7%; Pred. No. 1,23e-19;
Matches 39; Conservative 34; Mismatches 58; Indels 5; Gaps 5;

```

Search completed: Tue Jan 18 13:20:53 2000
Job time : 24 secs.

 WIDEOR (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

MSrch_PP protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Jan 18 13:19:47 2000; Maspar time 9.89 Seconds
 Tabular output not generated. 704.691 Million cell updates/sec

Title: >US-09-246-129-2
 Description: (1-174) from US09246129A.pep
 Perfect Score: 1314
 Sequence: 1 MRRFLSKVYSFPMRLILFL.....DISLVDYTKEDKTFGAFLL 174

Scoring table: PAM 150
 Gap 11

Searched: 122810 seqs, 40068593 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: pir60
 1:pir1 2:pir2 3:pir3 4:pir4

Statistics: Mean 44.271; Variance 85.869; scale 0.516

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	242	18.4	202	2	JN0869 tumor necrosis factor	8.30e-25
2	241	18.3	202	2	B27303 tumor necrosis factor	1.27e-24
3	232	17.7	279	2	A53062 Fas ligand - mouse	5.78e-23
4	229	17.4	197	2	JH0309 tumor necrosis factor	2.05e-22
5	229	17.4	204	2	S24641 lymphotoxin - bovine	2.05e-22
6	220	16.7	281	2	I38707 Fas ligand - human	8.93e-21
7	217	16.5	204	2	S17289 tumor necrosis factor	3.12e-20
8	212	16.1	278	2	A49266 fas ligand - rat	2.48e-19
9	200	15.2	205	1	OWHUX lymphotoxin alpha pre	3.43e-17
10	196	14.9	193	2	S06192 tumor necrosis factor	1.75e-16
11	196	14.9	232	2	S12606 tumor necrosis factor	1.75e-16
12	194	14.8	233	1	OWHUN tumor necrosis factor	3.93e-16
13	191	14.5	233	1	S22052 tumor necrosis factor	1.32e-15
14	191	14.5	306	2	I49139 lymphotoxin-beta - mo	1.32e-15
15	189	14.4	233	2	S13114 tumor necrosis factor	2.94e-15
16	189	14.4	234	2	JH0529 tumor necrosis factor	2.94e-15
17	189	14.4	235	1	OWMSN tumor necrosis factor	2.94e-15
18	187	14.2	234	2	QJ1344 tumor necrosis factor	6.57e-15
19	187	14.2	235	2	I54490 tumor necrosis factor	7.17e-14
20	181	13.8	234	2	A25451 tumor necrosis factor	5.17e-13
21	176	13.4	235	2	JU0029 tumor necrosis factor	7.66e-13
22	175	13.3	185	2	S52715 tumor necrosis factor	7.66e-13
23	175	13.3	235	2	S24642 tumor necrosis factor	7.66e-13

24	169	12.9	233	2	S11688 tumor necrosis factor	7.97e-12
25	138	10.5	244	2	A46066 lymphotoxin beta - hu	9.00e-07
26	112	8.5	307	2	S30432 hypothetical protein	6.94e-03
27	101	7.7	384	2	S03815 probable integrase -	2.25e-01
28	100	7.6	261	2	I53476 CD40 ligand - human	3.05e-01
29	100	7.6	455	2	D69785 beta-glucosidase homo	3.05e-01
30	98	7.5	532	2	A64796 signal-transducing hl	5.57e-01
31	97	7.4	261	2	S53090 CD40 ligand - bovine	7.52e-01
32	96	7.3	528	2	T00951 hypothetical protein	1.01e+00
33	96	7.3	1022	2	I53078 homeotic gene regulat	1.01e+00
34	95	7.3	1647	2	S45252 SNFbeta protein - hu	1.01e+00
35	95	7.2	229	2	E70405 ABC transporter - Aqu	1.36e+00
36	95	7.2	260	2	S21738 CD40 ligand - mouse	1.36e+00
37	94	7.2	358	2	F69809 spore germination pro	1.82e+00
38	93	7.1	3972	2	S75251 hypothetical protein	2.43e+00
39	92	7.0	527	2	A38631 transforming protein	3.24e+00
40	92	7.0	822	2	S30198 Na+/H+-exchanging pro	3.24e+00
41	91	6.9	245	2	S48272 hypothetical protein	4.31e+00
42	91	6.9	511	2	S55198 sulfate adenylyltrans	4.31e+00
43	91	6.9	551	2	A55582 cytochrome-c oxidase	4.31e+00
44	91	6.9	745	2	C70848 probable ioc2 protein	4.31e+00
45	91	6.9	1082	2	H70360 cation efflux system	4.31e+00

ALIGNMENTS

RESULT 1
 ENTRY JN0869 #type complete
 TITLE tumor necrosis factor beta - rat
 ORGANISM Rattus norvegicus #common_name Norway rat
 DATE 03-May-1994 #sequence_revision 03-May-1994 #text_change 08-Sep-1997

ACCESSIONS JN0869
 REFERENCE JN0868
 #authors Kwon, J.; Chung, I.Y.; Benveniste, E.N.
 #journal Gene (1993) 132:227-236
 #title Cloning and sequence analysis of the rat tumor necrosis factor-encoding genes.
 #cross-references M01D:94040766
 #accession JN0869

##molecule-type DNA
 ##residues 1-202 #label KMO
 ##cross-references GB:L00981; NID:q205253; PID:q205255
 ##note the authors translated codon CTC for residue 172 as ile
 COMMENT This protein is structurally related pleiotropic cytokine with
 overlapping biological functions, including an involvement in
 inflammatory processes.

GENETICS

#gene TNF-beta
 #introns 32/3; 66/1
 CLASSIFICATION #superfamily tumor necrosis factor
 KEYWORDS tumor
 SUMMARY #length 202 #molecular-weight 22111 #checksum 3095

Query Match 18.4%; Score 242; DB 2; Length 202;
 Best Local Similarity 36.5%; Pred. No. 8.30e-25;
 Matches 50; Conservative 22; Mismatches 57; Indels 8; Gaps 8;

Db	73	SLIMRANTDRAFLRHFSLNNSLLIPISGLFYVSQVFSG-ES-CSP-R-A-IPT-PI 126
Qy	39	ALMHEHLEGLAFTKRMNTNFLLIPESGDFFIYSQVFRMTSECSIRQAGRPRKFD 98
Db	127	YLAHEVQLPSSQYPPFVPLISAKRSYPPLOGPWRSMQGVFLSKDQOSTDTGDS 186
Qy	99	SIIVVITKVTDSIPETQLMGTKSYC-EVGSNMFQPIYLGMFSLQEGDKLNVVSDIS 157
Db	187	HLHFS-PSTVFFGAFAL 202
Qy	158	LVDTYTKEDKTFGAFLL 174
RESULT ENTRY	2	B27303 #type complete

TITLE tumor necrosis factor beta precursor - mouse
 ALTERNATE_NAMES lymphotoxin; TNF beta
 ORGANISM #formal_name Mus musculus #common_name house mouse
 DATE 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 08-Sep-1997

ACCESSIONS B27303; S01342; S10083; I56004; I48853; I55980
 REFERENCE A53679
 #authors Semon, D.; Kawashima, E.; Jongeneel, C.V.; Shakhov, A.N.;

#journal Nedospasov, S.A.
 #title Nucleic Acids Res. (1987) 15:9083-9084
 Nucleotide sequence of the murine TNF locus, including the TNF-alpha-(tumor necrosis factor) and TNF-beta-(lymphotoxin) genes.

#cross-references M01D:88067722
 #accession B27303
 #molecule_type DNA
 #residues 1-202 ##label SEM

REFERENCE S01342
 #authors Nedospasov, S.A.; Hilt, B.; Shakhov, A.N.; Dobrynin, V.N.;

#journal Kawashima, E.; Accolla, R.S.; Jongeneel, C.V.
 #title Nucleic Acids Res. (1986) 14:7713-7725
 The genes for tumor necrosis factor (TNF-alpha) and lymphotoxin (TNF-beta) are tandemly arranged on chromosome 17 of the mouse.

#cross-references M01D:87040736
 #accession S01342
 #molecule_type DNA
 #residues 1-11;139-160,'CG',163-178 ##label NED

REFERENCE S10083
 #authors Well, D.; Dautry, F.
 #journal Oncogene Res. (1988) 3:409-414
 Induction of tumor necrosis factor-alpha and -beta and interferon-gamma mRNA by interleukin 2 in murine lymphocytic cell lines.

#cross-references M01D:89144562
 #accession S10083
 #molecule_type mRNA
 #residues 6-202 ##label WEI

REFERENCE I56004
 #cross-references EMBL:X14800; NID:g54833; PID:g736269
 #accession I56004
 #authors Gardner, S.M.; Mock, B.A.; Hilgers, J.; Huppl, K.E.; Roeder, W.D.

#journal J. Immunol. (1987) 139:476-483
 #title Mouse lymphotoxin and tumor necrosis factor: Structural analysis of the cloned genes, physical linkage, and chromosomal position.

#cross-references M01D:87252204
 #accession I56004
 #status preliminary; translated from GB/EMBL/DBJ

REFERENCE I48853
 #cross-references GB:M17015; NID:g198860; PID:g387407
 #residues 1-25,'P',27-202 ##label RES
 #molecule_type DNA
 #authors Gray, P.W.; Chen, E.; Li, C.B.; Tang, W.T.; Ruddle, N.
 #journal Nucleic Acids Res. (1987) 15:3937
 The murine tumor necrosis factor-beta (lymphotoxin) gene sequence.

#cross-references M01D:87231097
 #accession I48853
 #status preliminary; translated from GB/EMBL/DBJ

REFERENCE I55980
 #molecule_type DNA
 #residues 1-202 ##label RE2
 #cross-references EMBL:X00137; NID:g54842; PID:g54843

REFERENCE I55980
 #authors Li, C.
 #journal J. Immunol. (1987) 138:4496-4501
 #title Cloning and expression of murine lymphotoxin cDNA.
 #cross-references M01D:87224127
 #accession I55980
 #status preliminary; translated from GB/EMBL/DBJ

#molecule_type mRNA
 #residues 1-202 ##label RE3

##cross-references GB:M16819; NID:g202088; PID:g202089
 COMMENT The first intron occurs in the 5'-untranslated region.
 GENETICS

#gene ToB
 #map_position 17
 #introns 32/3; 66/1
 CLASSIFICATION #superfamily tumor necrosis factor
 KEYWORDS cytokine; cytolysin; glycoprotein; lymphokine; macrophage
 SUMMARY #length 202 #molecular-weight 21998 #checksum 2011

Query Match 18.3%; Score 241; DB 2; length 202;
 Best local Similarity 36.5%; Pred. No. 1,27e-24;
 Matches 50; Conservative 22; Mismatches 57; Indels 8; Gaps 8;

Db 73 SLWRASDRAFLRHGFSUNSLIPTSGLYPVSQVFSG-ES-CSP-R-A-IPT-PI 126
 QY 39 ALHWEHLGLAFTKRNNTNKLFLPESGDITTYISQVFRGTSCESTRAGRPNKD 98

Db 127 YLAHEVQLFSSQYPFVPLLSAQSRYPGIQGPRSMYQGAFFLLSKDQSLHTDGIS 186
 QY 99 SITVITIKVTDYDEPTQLMGTKSYC-EVGSWMFQPIYLGAMFSLQEGDKLMVNSDIS 157

Db 187 HLHFS-PSSVFFGAFAL 202
 QY 158 LVDTYTKEDKTFEGAFLL 174

RESULT 3
 ENTRY A53062 #type complete
 TITLE Fas ligand - mouse
 ORGANISM #formal_name Mus musculus #common_name house mouse
 DATE 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 10-Sep-1997

ACCESSIONS A53062
 REFERENCE A53062
 #authors Takahashi, T.; Tanaka, M.; Brannan, C.I.; Jenkins, N.A.;

#journal Copeland, N.G.; Suda, T.; Nagata, S.
 #title Generalized lymphoproliferative disease in mice, caused by a point mutation in the Fas ligand.

#cross-references M01D:94185175
 #accession A53062
 #status preliminary

##molecule_type mRNA
 #residues 1-279 ##label TAK
 ##cross-references GB:U06948; NID:g473564; PID:g473565

SUMMARY #length 279 #molecular-weight 31442 #checksum 3291

Query Match 17.7%; Score 232; DB 2; length 279;
 Best local Similarity 35.3%; Pred. No. 5.78e-23;
 Matches 54; Conservative 32; Mismatches 50; Indels 17; Gaps 13;

Db 143 SVALITGNPMSRS-IP-LEWEDTYGTALISG-VKRRKGLVINENGLYFVSQVFRGO- 198
 QY 23 PVVAGQPTQHFKNQFPLHWEHLGLAFTKRNNTNKLFLPESGDITTYISQVFRGMT 82

Db 199 S-CNN--Q--P--LNH-KVYM-R-NSKYPEDVLVLEERLNTCTGTQIWAHSSYGAVF 247
 QY 83 SECEIRQAGRPKNKPSDITVITIKVDSYEPQQLMGTK-SVCEGSMWFQPIYLGAMF 141

Db 248 NLTSADHLVYNISQLSINF-EESKTFEGLYKL 279
 QY 142 SLQEGDKLMVNSDISLVDTYTKEDKTFEGAFLL 174

RESULT 4
 ENTRY JH0309 #type complete
 TITLE tumor necrosis factor beta precursor - rabbit
 ALTERNATE_NAMES lymphotoxin; TNF beta
 ORGANISM #formal_name Oryctolagus cuniculus #common_name domestic rabbit
 DATE 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 29-Jan-1999


```

#cross-references MUID:91139175
#accession 154482
#status translation not shown: translated from GB/EMBL/DBJ
##molecule_type DNA
##residues 1-124,'P',126-205 ##label RES
##cross-references GB:M55913; NID:g339742; PID:g339743
##experimental_source ancestral haplotype 57.1
#note 59-Asn was also found (ancestral haplotype 8.1)
REFERENCE
#authors A93350
#journal Gray, P.W.; Aggarwal, B.B.; Benton, C.V.; Bringman, T.S.;
#title Henzel, W.J.; Jarrett, J.A.; Leung, D.W.; Moffatt, B.; Ng,
#title P.; Sveinbjorn, L.P.; Palladino, M.A.; Nedwin, G.E.
#title Nature (1984) 312:721-724
#title Cloning and expression of cDNA for human lymphotoxin, a
#title lymphokine with tumour necrosis activity.
#cross-references MUID:85086243
#accession A93350
##molecule_type mRNA
##residues 1-205 ##label GRA
##cross-references GB:X01393; NID:g33444; PID:g33445
##experimental_source lymphoblastoid cell line RPMI-1788
REFERENCE
#authors A32877
#journal Goeddel, G.V.; Aggarwal, B.B.; Gray, P.W.; Leung, D.W.;
#title Nedwin, G.E.; Palladino, M.A.; Patton, J.S.; Penica, D.;
#title Shepard, H.M.; Sugarman, B.J.; Wong, G.H.W.
#title Cold Spring Harb. Symp. Quant. Biol. (1986) 51:597-609
#title Tumor necrosis factors: gene structure and biological
#title activities.
#cross-references MUID:87217059
#accession B32877
#status preliminary: not compared with conceptual translation
##molecule_type mRNA
##residues 35-205 ##label COE
REFERENCE
#authors A91906
#journal Kobayashi, Y.; Miyamoto, D.; Asada, M.; Obinata, M.; Osawa,
#title T.
#title J. Biochem. (1986) 100:727-733
#title Cloning and expression of human lymphotoxin mRNA derived from
#title a human T cell hybridoma.
#cross-references MUID:87057135
#accession A91906
##molecule_type mRNA
##residues 1-59,'N',61-205 ##label KOB
##cross-references GB:D00102; NID:g219913; PID:d1000508; PID:g219914
#note the authors translated the codon TAT for residue 156 as
#note Thr and ACC for residue 176 as Tyr
REFERENCE
#authors A61478
#journal Fukuda, S.; Ando, S.; Sanou, O.; Tanish, M.; Fujii, M.;
#title Masaki, N.; Nakamura, K.I.; Ando, O.; Toriigoe, K.;
#title Sugimoto, T.; Kurimoto, M.
#title Lymphokine Res. (1988) 7:175-185
#title Simultaneous production of natural human tumor necrosis
#title factor-alpha, -beta and interferon-alpha from BALL-1 cells
#title stimulated by HVJ.
#accession A61478
##molecule_type protein
##residues 56-79;86-95,'X',97,'X',99;119-151,'XX',154-162,'X',164,
#title 'X',166,'X',168,'X',170,'X',172-173 ##label FUK
REFERENCE
#accession S26951
#journal Voligt, C.G.; Maurer-Fogy, I.; Adolf, G.R.
#title FEBS Lett. (1992) 314:85-88
#title Natural human tumor necrosis factor beta (lymphotoxin).
#title Variable O-glycosylation at Thr(7), proteolytic processing,
#title and allelic variation.
#cross-references MUID:93083656
#accession S26951
##molecule_type protein
##residues 35-59,'N',61-205 ##label VOI
#note 60-Thr was also found
REFERENCE
#authors S34742
#journal Fukushima, K.; Watanabe, H.; Takeo, K.; Nomura, M.; Asahi,
#title T.; Yamashita, K.
#title Arch. Biochem. Biophys. (1993) 304:144-153

```

#title	N-linked sugar chain structure of recombinant human lymphotoxin produced by CHO cells: the functional role of carbohydrate as to its lectin-like character and clearance velocity.	
#cross-references	MUID:93311995	
#contents	annotation	
COMMENT	Secreted from mitogen-activated lymphocytes within 1-2 days after induction, this lymphotoxin can cause cytolytic of certain tumor cell lines and have an antiproliferative effect on others while having no detrimental effect on normal cells. It can also act synergistically with interferon gamma to kill certain transformed cell lines.	
COMMENT	This protein and TNF-alpha (tumor necrosis factor) are the products of different genes closely linked on chromosome 6. They are induced by similar mitogenic stimuli and have similar biological activities but are produced by different cell types and have different induction kinetics.	
GENETICS		
#gene	GDB:ITR; IT; TNFR	
#cross-references	GDB:120442; OMIM:153440	
#map-position	6p21.3-6p21.3	
#introns	33/3; 69/1	
#note	the first intron occurs before the initiator codon	
CLASSIFICATION	#superfamily tumor necrosis factor	
KEYWORDS	cytokine; cytotoxin; glycoprotein; homotrimer; lymphokine; macrophage	
FEATURE		
1-34	#domain signal sequence	#status predicted #label SIG
35-205	#product lymphotoxin	#status predicted #label MAT
41	#binding-site carbohydrate (Thr)	(covalent) (partial)
96	#status experimental\	
	#binding-site carbohydrate (Asn)	(covalent) #status experimental
SUMMARY	#length 205	#molecular-weight 22297 #checksum 9579
Query Match	15.2%;	Score 200; DB 1; Length 205;
Best Local Similarity	31.4%;	Pred. No. 3,43e-17;
Matches	43; Conservative	28; Mismatches 58; Indels 8; Gaps 7;
Db	76 SLIRANTDPAFLDQDGSFSSNLSLPTSGIFEVYSQVFSKKA--YGP-R-AT-SS-PL 129	
Qy	39 ALHHEHLGLAFIRKNRNTYTKRLFLPEBGDFYISQYFRMTSECEIRIQAGPKNPD 98	
Db	130 YLAHEVOLFSQYPEFHVPLISSOKMYVPEGLDPEWILSHMYGGAFLITGGDLSTHTDGP 189	
Qy	99 SITVIVIRKVDSPDEPQLIMGTKSYC-EVGSNMVFQPIYLGAMFSLQGDRLMVVSDIS 157	
Db	190 HLIVLS-PSYVFFGAFAL 205	
Qy	158 LVDTYKEDKTFFGAFLL 174	
RESULT	10	
ENTRY	S06192	#type fragment
TITLE	tumor necrosis factor alpha precursor - goat (fragment)	
ALTERNATE_NAMES	cachectin; TNF alpha	
ORGANISM	#formal_name Capra aegagrus hircus #common_name domestic goat	
DATE	28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 08-Sep-1997	
ACCESSIONS	S06192; S41867	
REFERENCE	S06192	
#authors	Goldstein, I.M.; Henner, D.; Talhouk, A.	
#submitted	submitted to the EMBL Data Library, March 1989	
#accession	S06192	
#molecule_type	mRNA	
#residues	1-193	#label GOT
REFERENCE	S41867	
#cross-references	EMBL:X14828; NID:g992; PID:g993	
REFERENCE	S41867	
#authors	Rimstad, E.	
#submitted	submitted to the EMBL Data Library, January 1994	
#accession	S41867	
#status	Preliminary	
#molecule_type	mRNA	

```

#cross-references M01D:50034181
#accession      I46659
#status
##molecule-type mRNA
##residues      44-232 ##label PAU
preliminary: translated from GB/EMBL/DBD

```

```
#title      Human tumour necrosis factor: precursor structure, expression
             and homology to lymphotoxin.
#cross-references MIMD:85086244
#accession  A93351
#molecule_type mRNA
```

 WISE (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

MSrch_PP protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Jan 18 13:18:22 2000; MasPar time 14.38 Seconds
 257.319 Million cell updates/sec

Tabular output not generated.

Title: >US-09-246-129-2
 Description: (1-174) from US09246129A.pep
 Perfect Score: 1314
 Sequence: 1 MRRFLSKVSPMRKLIPL.....DISLVDYTKEDKTFGAFLL 174

Scoring table: PAM 150
 Gap 11

Searched: 170751 seqs, 21266608 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: a-geneseq35
 1:part1 2:part2 3:part3 4:part4 5:parts 6:part6 7:part7
 8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
 14:part14 15:part15 16:part16 17:part17 18:part18
 19:part19 20:part20 21:part21 22:part22 23:part23
 24:part24 25:part25 26:part26 27:part27 28:part28
 29:part29 30:part30 31:part31 32:part32 33:part33
 34:part34 35:part35 36:part36 37:part37 38:part38
 39:part39

Statistics: Mean 31.581; Variance 131.813; scale 0.240

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	1301	99.0	174	18	R99453 Tumour necrosis facto	8.21e-118
2	252	19.2	169	25	W25079 Tumour metastasis sup	2.66e-13
3	247	18.8	180	23	W10875 FasL/alpha hybrid	7.76e-13
4	243	18.5	200	3	P50056 Sequence of murine ly	1.82e-12
5	241	18.3	169	8	R41866 Murine TNF-beta m	2.79e-12
6	238	18.1	157	2	R06359 TNF analogue	5.28e-12
7	237	18.0	239	30	W40364 Rat Fas ligand-like p	6.53e-12
8	235	17.9	291	31	W56762 Murine TRAIL polypept	9.98e-12
9	235	17.9	291	23	W19788 Mouse apoptosis induc	9.98e-12
10	235	17.9	291	29	W44353 Murine Agp-1	9.98e-12
11	232	17.7	137	14	R79067 Mouse Fas ligand (par	1.88e-11
12	232	17.7	138	14	R79068 Mouse Fas ligand (par	1.88e-11
13	232	17.7	179	14	R79069 Mouse Fas ligand (par	1.88e-11
14	232	17.7	279	14	R79098 Mouse Fas ligand	1.88e-11
15	231	17.6	279	36	W76332 Human TL2 (TRAIL), li	2.33e-11
16	231	17.6	279	18	R88357 Mouse Fas ligand	2.33e-11

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
17	231	17.6	281	27	W27134 Human Apoptosis induc	2.33e-11
18	231	17.6	281	23	W19777 Novel cytokine Apo-2	2.33e-11
19	231	17.6	281	31	W19787 Human apoptosis induc	2.33e-11
20	231	17.6	281	31	W56760 Human TRAIL polypept	2.33e-11
21	231	17.6	281	37	W76829 Human TR2 protein	2.33e-11
22	231	17.6	281	29	W44354 Human Agp-1	2.33e-11
23	228	17.4	157	23	W10877 TNF/LR-alpha hybrid	4.39e-11
24	228	17.4	171	8	R41864 Bovine TNF-beta m	3.56e-11
25	229	17.4	205	4	R23666 Ovine TNF-beta	3.56e-11
26	225	17.1	239	30	W40363 Mouse Fas ligand-like	8.28e-11
27	220	16.7	130	14	R79096 Human Fas ligand C-te	2.37e-10
28	220	16.7	136	33	W48955 Fas ligand fragment	2.37e-10
29	220	16.7	136	34	R88307 Human Fas ligand (par	2.37e-10
30	220	16.7	137	14	R79100 Human Fas ligand (par	2.37e-10
31	220	16.7	138	14	R88305 N-terminally deleted	2.37e-10
32	220	16.7	139	14	R79064 Human Fas ligand (par	2.37e-10
33	220	16.7	139	14	R88304 N-terminally deleted	2.37e-10
34	220	16.7	141	14	R88303 N-terminally deleted	2.37e-10
35	220	16.7	151	26	W16667 Human Fas ligand (hfa	2.37e-10
36	220	16.7	179	14	R79099 Human Fas ligand (par	2.37e-10
37	220	16.7	265	33	W48954 Non-cleavable Fas lig	2.37e-10
38	220	16.7	268	33	W48953 Non-cleavable Fas lig	2.37e-10
39	220	16.7	281	18	R98104 Human Fas ligand derl	2.37e-10
40	220	16.7	281	26	W27143 Human Fas ligand	2.37e-10
41	220	16.7	281	18	R88356 Human Fas ligand	2.37e-10
42	220	16.7	281	14	R79097 Human Fas ligand	2.37e-10
43	220	16.7	281	14	R77281 Human Fas-L protein	2.37e-10
44	220	16.7	281	36	W75959 Human Fas ligand	2.37e-10
45	220	16.7	287	26	W27146 CH3/Fas ligand domain	2.37e-10

ALIGNMENTS

RESULT 1
 ID R99453 standard; Protein: 174 AA.
 AC R99453;
 DT 14-SEP-1996 (first entry)
 DE Tumour necrosis factor-gamma.
 KW Tumour necrosis factor-gamma; TNF-gamma; cancer; tumour; gene therapy;
 KM wound healing.
 OS Homo sapiens.
 FH Key
 FT peptide
 FT Location/Qualifiers
 FT /label= Sig_peptide
 FT 26..174
 FT protein /label= Mat_protein
 FT
 PN W09614328-A1.
 PD 17-MAY-1996.
 PF 07-NOV-1994; U12880.
 PR 07-NOV-1994; W0-U12880.
 PR 21-NOV-1994; ZA-009229.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Ni J, Rosen CA, Yu G;
 PI WPI; 96-321550/32.
 DR N-PSDB; T30390.
 PT Tumour necrosis factor gamma and corresponding DNA and RNA - for
 PT treating cancer, facilitating wound healing and providing resistance
 PT against infection
 PS Claim 1, Fig 1; 69pp; English.
 CC Human tumour necrosis factor-gamma (R99453) is a novel member of
 CC the TNF family. It is encoded by a cDNA clone (T30390) isolated
 CC from a human umbilical vein endothelial cell cDNA library.
 CC Recombinant TNF-gamma can be expressed in bacterial, insect or
 CC mammalian cells transformed with a vector carrying the TNF-gamma
 CC coding sequence. It is used to inhibit tumour cell growth, to
 CC provide resistance to infection, to induce inflammatory activity,
 CC for wound healing etc., and to screen for (ant)agonists or to
 CC raise antibodies.
 SQ Sequence 174 AA.
 Query Match 99.0%; Score 1301; DB 18; Length 174;
 Best Local Similarity 99.4%; Pred. No. 8.21e-118;
 Matches 173; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 mtrfiskvysfmrklilflvfrvqrptqfhknqfapalhwelajlaftkrmnylnk 60
 |||||||
 QY 1 MRRFLSKVYSPFMRKLILFLVFPVVRQPTQFHKNQFAPALHWEHLGAFTRKRMNYLNK 60
 |||||||
 Db 61 flilpesdyfilyqvtfgmtsecelrgegrphkpsitvltkvdsipeptqlmg 120
 |||||||
 QY 61 flilpesdyfilyqvtfgmtsecelrgegrphkpsitvltkvdsipeptqlmg 120
 |||||||
 Db 121 tksvcegsnwfiplylgamfslqegdklmvnsdislvdytkedktffgafl 174
 |||||||
 QY 121 tksvcegsnwfiplylgamfslqegdklmvnsdislvdytkedktffgafl 174
 |||||||
 RESULT 2
 ID W25079 standard; Protein: 169 AA.
 AC W25079;
 DT 17-SEP-1998 (first entry)
 DE Tumour metastasis suppressor (TNF-alpha fusion protein).
 KW TNF-alpha; tumour metastasis; suppressor; tumour necrosis factor-alpha;
 KM cancer; treatment.
 OS Synthetic.
 FH Key
 FT Peptide 1..18 Location/Qualifiers
 FT /note="a preferred X"
 FT 19..169
 FT Protein
 PN CN101572-A.
 PD 19-APR-1995.
 PF 15-OCT-1993: 115007.
 PR 15-OCT-1993: CN-115007.
 PA (SOMA/) SOMA G.
 PI Mizuno D, Soma G;
 DR WPI: 97-260000/24.
 PT Tumour metastasis-suppressing agent - especially for use in
 PT veterinary medicine
 PS Disclosure: Page -: 11pp; Chinese
 CC This sequence comprises a new peptide inhibitor of cancer metastasis in
 CC animals. The inhibitor has a generic sequence: X'-X (I); (the amino acid
 CC sequence of the fourth circumscribed portion of TNF-alpha), where X =
 CC hydrogen or a polypeptide of any size; and X' = a polypeptide of 1-39
 CC residues, preferably as above, with the proviso that X-X' is not
 CC Met-Val-Arg-Ser-Ser-Ser-Arg-Thr-Pro-Ser-Asp-Lys-Pro-Val-Ala-His-Val-Val.
 SC Sequence 169 AA.
 Query Match 19.2%; Score 252; DB 25; Length 169;
 Best Local Similarity 30.8%; Pred. No. 2,66e-13;
 Matches 44; Conservative 37; Mismatches 54; Indels 8; Gaps 4;
 Db 27 lqylrranallangvelrdnglvpvpsqglylysqvfkqggqgcpsthlthpschv 86
 |||||||
 QY 40 LHMHEHLGIAFTKRNMYTNKFLIPESGDYFIYSQVFRGMTSCSEIR-QAGRPNKPD 98
 |||||||
 Db 87 lltthstiaavgyctcnllsaikspcqrpepgeakpwyepilqgvfqlqgdrtisa 146
 |||||||
 QY 99 SITVITKVTDSYPTQLMGTKSYC--EV--GSN---WFOPIYLGAMFSLQEGDKLMV 151
 |||||||
 Db 147 ealnrdpyldfaesggyvfigtal 169
 |||||||
 QY 152 NVSDISLVDTYTKEDKTFEGAFLL 174
 |||||||
 RESULT 3
 ID W10875 standard; Protein: 180 AA.
 AC W10875;
 DT 25-SEP-1997 (first entry)
 DE Fast/LT-alpha hybrid.
 KW C-terminal extracellular domain; CD40 ligand; mutated; trimerisation;
 KM lymphotoxin; LT; beta; alpha; tumour necrosis factor; TNF; inhibitor;
 KW signalling; receptor; Fas; treating; autoimmune disease; diabetes;
 KW rheumatoid arthritis; inflammatory bowel disease; HIV proliferation;
 OS immune response; tissue graft.
 FH Synthetic.
 FH Key Location/Qualifiers

FT misc_difference 44
 FT /note="LT-alpha subunit association domain residue"
 FT misc_difference 68..70
 FT /note="LT-alpha subunit association domain residues"
 FT misc_difference 70
 FT /note="given as O in the specification"
 FT misc_difference 92
 FT /note="LT-alpha subunit association domain residue"
 FT misc_difference 94
 FT /note="LT-alpha subunit association domain residue"
 FT misc_difference 122..127
 FT /note="LT-alpha subunit association domain residues"
 FT misc_difference 126
 FT /note="given as O in the specification"
 FT misc_difference 129..130
 FT /note="LT-alpha subunit association domain residues"
 FT misc_difference 139
 FT /note="LT-alpha subunit association domain residue"
 FT misc_difference 146
 FT /note="LT-alpha subunit association domain residue"
 FT misc_difference 172
 FT /note="LT-alpha subunit association domain residue"
 FT misc_difference 177..178
 FT /note="LT-alpha subunit association domain residues"
 FT W09640774-A1.
 PN 19-DEC-1996.
 PD 06-JUN-1996: 009773.
 PF 07-JUN-1995: US-476074.
 PR (BioJ) BIOGEN INC.
 PA Browning JL, Karpusas MN, Meier W;
 DR WPI: 97-087055/08.
 PT Heteromeric complexes comprising lymphotoxin sub-units - useful as
 PT inhibitors of signalling by TNF- and TNF-related receptors, for
 PT treating autoimmune disease, graft rejection. etc.
 PS Disclosure: Fig 10: 87pp; English.
 CC This sequence comprises the C-terminal extracellular domain of the Fas
 CC ligand altered to drive trimerisation with wild-type lymphotoxin (LT)-
 CC beta. LT subunit association domains potentiate subunit associations into
 CC an active trimeric ligand. Tumour necrosis factor (TNF)-related ligand
 CC monomers mutated in their respective subunit association domains can
 CC form heteromeric complexes with LT subunits. LT-alpha/beta heteromeric
 CC complexes and LT heterotrimeric complexes are useful as inhibitors of
 CC signalling by TNF receptors (R) and TNF-related receptors, such as CD40,
 CC Fas and LT-beta receptors. By inhibiting TNF-R signalling the complexes
 CC (esp., LT-alpha2/beta1) can be used for treating autoimmune diseases
 CC including diabetes, rheumatoid arthritis, inflammatory bowel disease and
 CC HIV proliferation. known to be improved by TNF inhibition. The complexes
 CC are also useful for preventing or reducing the severity of an immune
 CC response to a tissue graft.
 SC Sequence 180 AA.
 Query Match 18.8%; Score 247; DB 23; Length 180;
 Best Local Similarity 39.1%; Pred. No. 7,76e-13;
 Matches 54; Conservative 25; Mismatches 43; Indels 16; Gaps 10;
 Db 57 plwedtygivilxg-vkykkgglvlnetgilyfvsxvfrqg-s-cnplps---hk-- 108
 |||||||
 QY 39 ALHMEHLGIAFTKRNMYTNKFLIPESGDYFIYSQVFRGMTSCSEIR-QAGRPNKPD 98
 |||||||
 Db 109 ---ym-rnskypqdlvllsskmycttgqmwahssylgaafnltsadllynnvsl 163
 |||||||
 QY 99 SITVITKVTDSYPTQLMGTKSYC--EV--GSN---WFOPIYLGAMFSLQEGDKLMV 156
 |||||||
 Db 164 slvnf-eestffgafl 180
 |||||||
 QY 157 SLVDYTKEDKTFEGAFLL 174
 |||||||
 RESULT 4
 ID P50056 standard; Protein: 200 AA.
 AC P50056;
 DT 04-SEP-1991 (first entry)
 DE Sequence of murine lymphotoxin.


```

QY      23 PVRQPTQHFKNQFALHWEHELGLAFTKRNKNTNKLIPESGDYFIYSQVTFKMT 82
      57 s-cnn--q--p--lnh-kvym-r-nskypedlvimeekrllyctgqlwshssylgavf 105
      83 SECSEIRQGRPNKPDSTIVITKVTDSYPEPTQLMGTR-SVCEVGSNMPPIYLGAMF 141
      106 nltsadhlyvnisqslinf-eesktffgllykl 137
      142 SLOEGDKLMVNSDISLVYTRKEDKTFEGAFLL 174

```

RESULT 12
ID R79068 standard; Protein: 138 AA.

```

AC R79068;
DE 22-FEB-1996 (first entry)
KW Fas ligand; Tumour Necrosis factor family; apoptosis; cell death;
KW Fas cell surface antigen; Fas-L; mouse.
OS Mus musculus.
PN MO9513293-A1.
PD 18-MAY-1995.
PR 10-NOV-1994; J01899.
PR 10-NOV-1993; JP-305975.
PR 13-DEC-1993; JP-342526.
PR 18-MAR-1994; JP-074344.
PR 08-JUL-1994; JP-180955.
PR 07-SEP-1994; JP-239363.
PR 18-OCT-1994; JP-278378.
PA (MOCH) MOCHIDA PHARM CO LTD.
PA (OSAB-) OSAKA BIOSCIENCE INST.
PI Nagata S, Nakamura N, Suda T, Takahashi T;
DR WPI: 95-194031/25.
DR N-PSDB: Q99497.
PT Peptide which binds to Fas antigen, and antibody reactive with it
PT for treatment and diagnosis of viral or auto-immune diseases
PS Claim 10: Page 221-222: 300pp: Japanese.
CC Fas ligands or active fragments able to induce apoptosis in cells
CC which express the Fas cell surface antigen are claimed. The
CC proteins are isolated from human, rat and mouse sources. The present
CC sequence represents part of the mouse Fas ligand.
SQ Sequence 138 AA;

```

Query Match 17.7%; Score 232; DB 14; Length 138;

Best Local Similarity 35.3%; Pred. No. 1.88e-11;

Matches 54; Conservative 32; Mismatches 50; Indels 17; Gaps 13;

```

Db      2 svahltgnphstrs-lp-lwedtygtalysg-vkykkgjlvineqlyfyvskvyffrg- 57
      23 PVRQPTQHFKNQFALHWEHELGLAFTKRNKNTNKLIPESGDYFIYSQVTFKMT 82
      58 s-cnn--q--p--lnh-kvym-r-nskypedlvimeekrllyctgqlwshssylgavf 106
      83 SECSEIRQGRPNKPDSTIVITKVTDSYPEPTQLMGTR-SVCEVGSNMPPIYLGAMF 141
      107 nltsadhlyvnisqslinf-eesktffgllykl 138
      142 SLOEGDKLMVNSDISLVYTRKEDKTFEGAFLL 174

```

RESULT 13

ID R79069 standard; Protein: 179 AA.

```

AC R79069;
DE 22-FEB-1996 (first entry)
KW Mouse Fas ligand (partial sequence).
KW Fas ligand; Tumour Necrosis factor family; apoptosis; cell death;
KW Fas cell surface antigen; Fas-L; mouse.
OS Mus musculus.
PN MO9513293-A1.
PD 18-MAY-1995.
PR 10-NOV-1994; J01899.
PR 10-NOV-1993; JP-305975.
PR 13-DEC-1993; JP-342526.

```

```

PR 18-MAR-1994; JP-074344.
PR 08-JUL-1994; JP-180955.
PR 07-SEP-1994; JP-239363.
PR 18-OCT-1994; JP-278378.
PA (MOCH) MOCHIDA PHARM CO LTD.
PA (OSAB-) OSAKA BIOSCIENCE INST.
PI Nagata S, Nakamura N, Suda T, Takahashi T;
DR WPI: 95-194031/25.
DR N-PSDB: Q99498.
PT Peptide which binds to Fas antigen, and antibody reactive with it
PT for treatment and diagnosis of viral or auto-immune diseases
PS Claim 11: Page 222-224: 300pp: Japanese.
CC Fas ligands or active fragments able to induce apoptosis in cells
CC which express the Fas cell surface antigen are claimed. The
CC proteins are isolated from human, rat and mouse sources. The present
CC sequence represents part of the mouse Fas ligand.
SQ Sequence 179 AA;

```

Query Match 17.7%; Score 232; DB 14; Length 179;

Best Local Similarity 35.3%; Pred. No. 1.88e-11;

Matches 54; Conservative 32; Mismatches 50; Indels 17; Gaps 13;

```

Db      43 svahltgnphstrs-lp-lwedtygtalysg-vkykkgjlvineqlyfyvskvyffrg- 98
      23 PVRQPTQHFKNQFALHWEHELGLAFTKRNKNTNKLIPESGDYFIYSQVTFKMT 82
      99 s-cnn--q--p--lnh-kvym-r-nskypedlvimeekrllyctgqlwshssylgavf 147
      83 SECSEIRQGRPNKPDSTIVITKVTDSYPEPTQLMGTR-SVCEVGSNMPPIYLGAMF 141
      148 nltsadhlyvnisqslinf-eesktffgllykl 179
      142 SLOEGDKLMVNSDISLVYTRKEDKTFEGAFLL 174

```

RESULT 14

ID R79068 standard; Protein: 279 AA.

AC R79068;

DE 21-FEB-1996 (first entry)

DE Mouse Fas ligand.

KW Fas ligand; Tumour Necrosis factor family; apoptosis; cell death;

KW Fas cell surface antigen; Fas-L; mouse.

OS Mus musculus.

PH Key Location/Qualifiers

FT region 25..78

FT /label= proline-rich

FT domain 79..100

FT /label= transmembrane-anchor

FT modified_site 117

FT /label= N-glycosylation_site

FT modified_site 131

FT /label= N-glycosylation_site

FT modified_site 182

FT /label= N-glycosylation_site

FT modified_site 248

FT /label= N-glycosylation_site

FT modified_site 258

FT /label= N-glycosylation_site

PN MO9513293-A1.

PD 18-MAY-1995.

PR 10-NOV-1994; J01899.

PR 10-NOV-1993; JP-305975.

PR 13-DEC-1993; JP-342526.

PR 18-MAR-1994; JP-074344.

PR 08-JUL-1994; JP-180955.

PR 07-SEP-1994; JP-239363.

PR 18-OCT-1994; JP-278378.

PA (MOCH) MOCHIDA PHARM CO LTD.

PA (OSAB-) OSAKA BIOSCIENCE INST.

PI Nagata S, Nakamura N, Suda T, Takahashi T;

DR WPI: 95-194031/25.

DR N-PSDB: Q99499.

PT Peptide which binds to Fas antigen, and antibody reactive with it
 PS for treatment and diagnosis of viral or autoimmune diseases
 CC Claim 12; Fig 23-24; 300pp; Japanese.
 CC A clone (094156) contg. an insert coding for the mouse Fas ligand
 CC was isolated using probes derived from a rat Fas-L clone (see
 CC 094153). The mouse sequence contains an open reading frame coding
 CC for a 279 amino acid sequence with 5 N-glycosylation sites and a
 CC proline-rich region. The deduced mouse amino acid sequence has
 CC 91.4% homology with rat Fas-L. Fas ligands are able to induce
 CC apoptosis in cells which express the Fas cell surface antigen.
 SQ Sequence 279 AA;

Query Match 17.7%; Score 232; DB 14; Length 279;
 Best Local Similarity 35.3%; Pred. No. 1.88e-11;
 Matches 54; Conservative 32; Mismatches 50; Indels 17; Gaps 13;

Db 143 svahltpmhrs-ip-lwedtygtalssg-vkykkgslvneqglyvyskyvfrg- 198

QY 23 PVRKOTPTOHKQPFALHHEHLGLAFKRNKNTNKLLEPSGDYFYSQVTFRGMT 82

Db 199 s-cnn--q--p--lhh-kvym-r-nskypedlvimeekrllyctgqiwahssylgavf 247

QY 83 SECSEIRQAGRNKRPDSIRIVITKVTDSYEPPTQLMGTRK-SVCEVGSNMFPYILGAMF 141

Db 248 nltssadhlvynisqslinf-eesktffglykl 279

QY 142 SLQEGDKLMVNVSDISLDVYTKEDKTFEGAFLL 174

RESULT 15

ID W76332 standard; Protein: 279 AA.
 AC W76332;

DT 11-JAN-1999 (first entry)
 DE Human TL2 (TRAIL), ligand for TR5.

KW TL2; TRAIL; tumour necrosis related receptor; TR5; inflammation;
 KW arthritis; septicaemia; transplant rejection; autoimmune disease;
 KW inflammatory bowel disease; graft versus host disease; infection;
 KW stroke; ischaemia; acute respiratory disease syndrome; psoriasis;
 KW restenosis; brain injury; AIDS; bone disease; cancer;
 KW atherosclerosis; Alzheimer's disease; human; therapy; diagnosis;
 KW ligand.

KM Homo sapiens.
 OS EP-867509-A2.

PN 30-SEP-1998.
 PD 04-FEB-1998; 300827.

PR 28-JUL-1997; US-901469.
 PR 05-FEB-1997; US-795910.

PA (SMK) SMITHKLINE BEECHAM CORP.
 PI Lyn SDP, Tan KB, Truneh A, Young PR;

DR WPI; 98-497862/43.
 PT New polynucleotide encoding TR5 polypeptide - used to diagnose,

PT prevent and treat e.g. inflammation, arthritis, septicaemia,
 PT autoimmune diseases, infections, stroke, ischaemia, ARDS, psoriasis,
 PT restenosis, brain injury, AIDS and bone diseases

PS Disclosure; Page 17-18; 22pp; English.
 CC This is the amino acid sequence of human TL2 (also known as TRAIL),

CC which has newly been discovered to be a ligand of human tumour
 CC necrosis related receptor TR5 (see W76331). This TR5 polypeptide
 CC of the invention and TL2 can be used in screening processes for

CC compounds which bind the receptor, or its ligand, and which
 CC activate (agonists) or inhibit (antagonists) the receptor or TL2.
 CC Treatment of a subject with the need to inhibit TR5 polypeptide,
 CC activity comprises administering an antagonist to the polypeptide,
 CC administering a nucleic acid that inhibits the expression of the
 CC nucleotide sequence encoding the polypeptide and/or administering a

CC polypeptide that competes with the polypeptide for its ligand,
 CC substrate or receptor. The active agents can be used for the
 CC treatment of chronic and acute inflammation, arthritis,

CC septicemia, autoimmune diseases (e.g. inflammatory bowel disease,
 CC psoriasis), transplant rejection, graft vs host disease, infection,
 CC stroke, ischaemia, acute respiratory disease syndrome, restenosis,
 CC brain injury, AIDS, bone diseases, cancer (e.g. lymphoproliferative

CC disorders), atherosclerosis and Alzheimer's disease.

SQ Sequence 279 AA;

Query Match 17.6%; Score 231; DB 36; Length 279;
 Best Local Similarity 34.6%; Pred. No. 2.33e-11;
 Matches 47; Conservative 33; Mismatches 44; Indels 12; Gaps 11;

Db 152 wessrsqnsflsn-lhlngevlhnekgyfyysqlyfr-fg-e--elkentkndk-qmv 205

QY 42 WEHEL-GLAFKRNKNTNKLLEPSGDYFYSQVTFRGMTSECEIRQAGRNKRPDSI 100

Db 206 qy-lykyt-sypdpdllmksarnscwskdaeylysiyggifefekendrffvstnehl 263

QY 101 TVVITKVTDSIPEPTQLMGTRKSV-EGVGSNM-FQPIILGAMFSLQEGDKLMVNVSDISL 158

Db 264 ld-mdeasffgaflv 278

QY 159 VDYTKEDKTFEGAFLL 174

Search completed: Tue Jan 18 13:19:30 2000
 Job time : 68 secs.

